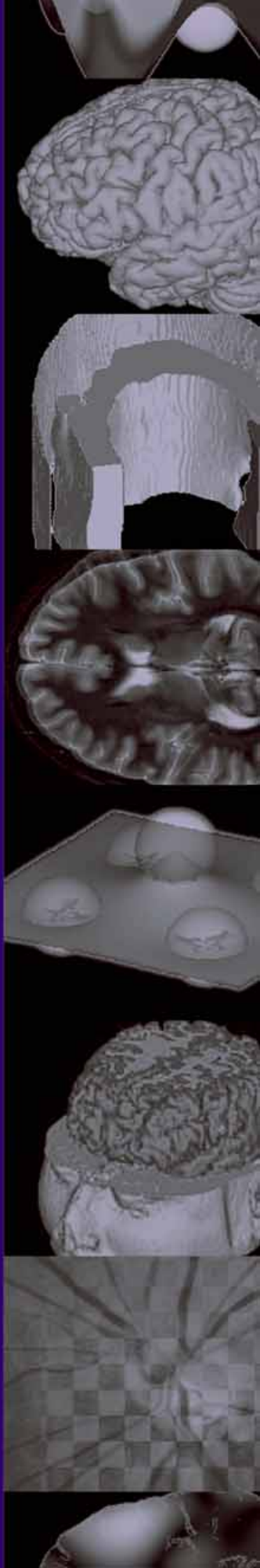
Handbook of Biomedical Image Analysis

Volume II:

Segmentation Models

Part B



Jasjit S. Suri
David L. Wilson
Swamy Laxminarayan

Handbook of Biomedical Image Analysis

TOPICS IN BIOMEDICAL ENGINEERING
INTERNATIONAL BOOK SERIES

Series Editor: Evangelia Micheli-Tzanakou
Rutgers University
Piscataway, New Jersey

Signals and Systems in Biomedical Engineering:
Signal Processing and Physiological Systems Modeling
Suresh R. Devasahayam

Models of the Visual System
Edited by George K. Hung and Kenneth J. Ciuffreda

PDE and Level Sets: Algorithmic Approaches to Static and Motion Imagery
Edited by Jasjit S. Suri and Swamy Laxminarayan

Frontiers in Biomedical Engineering:
Proceedings of the World Congress for Chinese Biomedical Engineers
Edited by Ned H.C. Hwang and Savio L-Y. Woo

Handbook of Biomedical Image Analysis:
Volume I: Segmentation Models Part A
Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

Handbook of Biomedical Image Analysis:
Volume II: Segmentation Models Part B
Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

Handbook of Biomedical Image Analysis:
Volume III: Registration Models
Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

---

Handbook of Biomedical Image Analysis

Volume II: Segmentation Models Part B

Edited by

Jasjit S. Suri
Department of Biomedical Engineering
Case Western Reserve University
Cleveland, Ohio

David L. Wilson
Department of Biomedical Engineering
Case Western Reserve University
Cleveland, Ohio

and

Swamy Laxminarayan
Institute of Rural Health
Idaho State University
Pocatello, Idaho

Kluwer Academic / Plenum Publishers
New York, Boston, Dordrecht, London, Moscow

ISBN 0-306-48605-9
eISBN 0-306-48606-7
set ISBN: 0-387-23126-9

233 Spring Street, New York, New York 10013

http://www.wkap.nl/

10 9 8 7 6 5 4 3 2 1

A C.I.P. record for this book is available from the Library of Congress

Printed in the United States of America

Jasjit Suri dedicates this handbook to
his youngest uncle Paramjeet Chadha and his immediate family:
his late sister Sharan, his late brother Amarjeet, and his
late parents Kulwant Kaur and Udam S. Chadha
(Fellow of Royal Institute of London).

David Wilson dedicates this handbook to his
family and students.

Swamy Laxminarayan dedicates
this book in memory of his beloved parents who were a
constant source of inspiration in his life and to his in-laws
Corie and Derk Zwakman for their genuine sense
of family attachments and friendship.

Contributors

Alejandro F. Frangi, Ph.D.
University of Zaragoza,
Zaragoza, Spain

Anand Manhoar, M.S.
University of South Florida,
Tampa, FL, USA

Bruno M. Carvalho, M.S.
University of Pennsylvania,
Philadelphia, USA

Chun Yuan, Ph.D.
University of Washington,
Seattle, WA, USA

David L. Wilson, Ph.D.
Case Western Reserve University,
Cleveland, OH, USA

Dirk Vandermeulen, Ph.D.
Katholieke Universiteit Leuven,
Leuven, Belgium

Emiliano D'Agostino, Ph.D.
Katholieke Universiteit Leuven,
Leuven, Belgium

Frederik Maes, Ph.D.
Katholieke Universiteit Leuven,
Leuven, Belgium

Gabor Szekely, Ph.D.
Swiss Federal Institute of Technology,
Zurich

Gabor T. Herman, Ph.D.
CUNY, New York, NY, USA

Jasjit S. Suri, Ph.D.
Case Western Reserve University,
Cleveland, OH, USA

Jean-Claude Klein, Ph.D.
Ecole des Mines de Paris,
Fontainebleau, France

Jeffrey Duerk, Ph.D.
Case Western Reserve University,
Cleveland, OH, USA

Jian Yang, Ph.D.
University of Zaragoza,
Zaragoza, Spain

John J. Heine, Ph.D.
University of South Florida,
Tampa, FL, USA

Jonathan Lewin, M.D.
Case Western Reserve University,
Cleveland, OH, USA

Keir Bovis, Ph.D.
University of Exeter,
Exeter, UK

Koen Van Leemput, Ph.D.
Helsinki University Central Hospital,
Helsinki, Finland

Koon-Pong Wong, Ph.D.
Hong Kong Polytechnic University,
Hung Hom, Kowloon Hong Kong

Maria Kallergi, Ph.D.
University of South Florida,
Tampa, FL, USA

Marla R. Hersh, M.D.
University of South Florida,
Tampa, FL, USA

Martín Laclaustra, M.D., Ph.D.
Aragon Institute of Health Sciences,
Zaragoza, Spain

Matthias Harders, Ph.D.
Swiss Federal Institue of Technology,
Zurich

Misael Rosales, Ph.D.
Universidad de los Andes,
Mérida/Venezuela

Mugdha Tembey, M.S.
University of South Florida,
Tampa, FL, USA

Oriol Pujol, Ph.D.
Universitat Autnoma de Barcelona,
Barcelona, Spain

Olivier Salvado, M.S.
Case Western Reserve University,
Cleveland, OH, USA

Paul Suetens, Ph.D.
Katholieke Universiteit Leuven,
Leuven, Belgium

Petia, Raveda, Ph.D.
Universitat Autnoma de Barcelona,
Barcelona, Spain

Philippe C. Cattin, Ph.D.
Swiss Federal Institute of Technology,
Zurich

Sameer Singh, Ph.D.
University of Exeter,
Exeter, UK

Sameer Singh, Ph.D.
University of Exeter,
Exeter, UK

Shuyu Yang, Ph.D.
Texas Tech University,
Lubbock, TX, USA

Siddharth Srivastava, Ph.D.
Katholieke Universiteit Leuven,
Leuven, Belgium

Sierra, Ph.D.
Swiss Federal Institute of Technology,
Zurich

Sunanda Mitra, Ph.D.
Texas Tech University,
Lubbock, TX, USA

Swamy Laxminarayan, D.Sc.
State University of Idaho,
Pocatello, ID, USA

Thomas Walter, Ph.D.
Ecole des Mines de Paris,
Fontainebleau, France

Vasanth Pappu, B.S.
Case Western Reserve University,
Cleveland, OH, USA

William S. Kerwin, Ph.D.
University of Washington,
Seattle, WA, USA

Yoshinobu Sato, Ph.D.
Osaka University,
Osaka, Japan

Zachary E. Miller, Ph.D.
University of Washington,
Seattle, WA, USA

Acknowledgments

This book is the result of collective endeavor from several noted engineering and computer scientists, mathematicians, medical doctors, physicists, and radiologists. The editors are indebted to all of their efforts and outstanding scientific contributions. The editors are particularly grateful to Drs. Petia Reveda, Alex Falco, Andrew Laine, David Breen, David Chopp, C. C. Lu, Gary Christensen, Dirk Vandermeulen, Aly Farag, Alejandro Frangi, Gilson Antonio Giraldi, Gabor Szekely, Pierre Hellier, Gabor Herman, Ardeshir Coshtasby, Jan Kybic, Jeff Weiss, Jean-Claude Klein, Majid Mirmehdi, Maria Kallergi, Yangming Zhu, Sunanda Mitra, Sameer Singh, Alessandro Sarti, Xioping Shen, Calvin R. Maurer, Jr., Yoshinobu Sato, Koon-Pong Wong, Avdhesh Sharma, Rakesh Sharma, and Chun Yuan and their team members for working with us so closely in meeting all of the deadlines of the book. We would like to express our appreciation to Kluwer Publishers for helping create this invitational handbook. We are particularly thankful to Aaron Johnson, the acquisition editor and Shoshana Sternlicht for their excellent coordination of the book at every stage.

Dr. Suri thanks Philips Medical Systems, Inc., for the MR datasets and encouragements during his experiments and research. Special thanks are due to Dr. Larry Kasuboski and Dr. Elaine Keeler from Philips Medical Systems, Inc., for their support and motivations. Thanks are also due to my past Ph.D. committee research professors, particularly Professors Linda Shapiro, Robert M. Haralick, Dean Lytle, and Arun Somani, for their encouragements.

We extend our appreciations to Drs. Ajit Singh, Siemens Medical Systems, George Thoma, chief, Imaging Science Division, National Institutes of Health, Dr. Sameer Singh, University of Exeter, UK, for his motivations.

Special thanks go to the book series editor, Professor Evangelia Micheli Tzanakou for advising us on all aspects of the book.

We thank the IEEE Press, Academic Press, Springer-Verlag Publishers, and several medical and engineering journals for permitting us to use some of the images previously published in these journals.

Finally, Jasjit Suri thanks his wife Malvika Suri for all the love and support she has showered over the years and to our baby Harman whose presence is always a constant source of pride and joy. I also express my gratitude to my father, a mathematician, who inspired me throughout my life and career, and to my late mother, who most unfortunately passed away a few days before my Ph.D. graduation, and who so much wanted to see me write this book. Special thanks to Pom Chadha and his family, who taught me life is not just books. He is my of my best friends. I would like to also thank my in-laws who have a special place for me in their hearts and have shown lots of love and care for me.

David Wilson acknowledges the support of the Department of Biomedical Engineering, Case Western Reserve University, in this endeavor. Special thanks are due to the many colleagues and students who make research in biomedical engineering an exciting, wondrous endeavor.

Swamy Laxminarayan expresses his loving acknowledgments to his wife Marijke and to his kids, Malini and Vinod, for always giving the strength of mind amidst all lifes frustrations. The book kindies fondest memories of my late parents who made many personal sacrifices that helped shape our careers and the support of my family members who were always there for me when I needed them most. I have shared many ideas and thoughts on the book with numerous of my friends and colleagues in the discipline. I acknowledge their friendship, feedbacks, and discussions with particular thanks to Professor David Kristol of the New Jersey Institute of Technology. Peter Brett of Ashton University, Ewart Carson of the City University, London, Laura Roa of the University of Sevilla in Spain, and Jean Louis Coatrieux of the University of Rennes in France for their constant support over the past two decades.

Preface

In Chapter 1 we present in detail a framework for fully automated brain tissue classification. The framework consists of a sequence of fully automated state of the art image registration (both rigid and nonrigid) and image segmentation algorithms. Models of the spatial distribution of brain tissues are combined with models of expected tissue intensities, including correction of MR bias fields and estimation of partial voluming. We also demonstrate how this framework can be applied in the presence of lesions.

Chapter 2 presents the intravascular ultrasound (IVUS), which is a tomographic imaging technique that has provided unique tool for observation and supervision of vessel structures and exact vascular dimensions. In this way, it has contributed to the better understanding of the coronary content and processes: vascular remodelling, plaque morphology, and evolution, etc. Most investigators are convinced that the best way to detect plaque ruptures is by IVUS sequences. At the same time, cardiologists confirm that due to the "speckle nature" of IVUS images, conventional IVUS imaging is difficult to clearly diagnose potentially vulnerable plaques due to the image resolution, lack of contours, speckle motion, etc. Advanced automatic classification techniques can significantly help the physicians take decisions about different classes of tissue morphology. The characterization of tissue and plaque involves different problems. Image feature space determines the reliable descriptions that should be sufficiently expressive to capture differences between different classes but at the same time should not increase unnecessarily the complexity of the classification problem. We consider and compare a wide set of different feature spaces (Gabor filters, DOG filters, cooccurrence matrices, binary local patterns, etc). In particular, we show that

the binary local patterns represent an optimal description of ultrasound regions that at the same time allow real-time processing of images. After reviewing the IVUS classification works available in the bibliography, we present a comparison between classical and advanced classification techniques (principal component analysis, linear discriminant analysis, nonparametric discriminant analysis, Kernel principal component analysis, Kernel fisher analysis, etc.). The classification "goodness" of IVUS regions can be significantly improved by applying multiple classifiers (boosting, adaboost, etc.). The result of the classification techniques represents a map of classified pixels that still need to be organized in regions. The technique of snakes (deformable models) is a convenient way to organize regions of pixels with similar characteristics. Incorporating the classification map or the likelihood map into the snake framework, allows to organize pixels into compact image regions representing different plaque zones of IVUS images.

Chapter 3 is dedicated to functional imaging techniques. The last few decades of the twentieth century have witnessed significant advances in multidimensional medical imaging, which enabled us to view noninvasively, the anatomic structure of internal organs with unprecedented precision and to recognize any gross pathology of organs and diseases without the need to "open" the body. This marked a new era of medical diagnostics with many invasive and potentially morbid procedures being substituted by noninvasive cross-sectional imaging. Continuing advances in instrumentation and computer technologies also accelerated the development of various multidimensional imaging modalities that possess a great potential for providing, in addition to structural information, dynamic, and functional information on biochemical and pathophysiologic processes or organs of the human body. There is no doubt that substantial progress has been achieved in delivering health care more efficiently and in improving disease management, and that diagnostic imaging techniques have played a decisive role in routine clinical practice in almost all disciplines of contemporary medicine. With further development of functional imaging techniques, in conjunction with continuing progress in molecular biology and functional genomics, it is anticipated that we will be able to visualize and determine the actual molecular errors in a specific disease very soon, and be able to incorporate this biological information into clinical management of that particular group of patients. This is definitely not achievable with the use of structural imaging techniques. In this chapter, we will take a quick tour of a functional imaging technique called

positron emission tomography (PET), which is a primer biologic imaging tool being able to provide in vivo quantitative functional information in most organ systems of the body. An overview of this imaging technique, including the basic principles and instrumentation, methods of image reconstruction from projections, some specific correction factors necessary to achieve quantitative images are presented. Basic assumptions and special requirements for quantitation are briefly discussed. Quantitative analysis techniques based on the framework of tracer kinetic modeling for absolute quantification of physiological parameters of interest are also introduced in this chapter.

Pancreatic cancer is a difficult to diagnose and lethal disease. In Chapter 4, we present the Helical computed tomography (CT), which is currently the imaging modality of choice for the detection, diagnosis, and evaluation of pancreatic tumors. Despite major technological advances, helical CT imaging still presents imaging limitations as well as significant challenges in the interpretation process. Computer methodologies could assist radiologists and oncologists in the interpretation of CT scans and improve the diagnosis and management of the patients with pancreatic cancer. However, few computer aided detection (CADetection) or diagnosis (CADiagnosis) techniques have been developed for pancreatic cancer and this area remains seriously understudied and unexplored. This chapter aims at introducing the problem of pancreatic cancer and the limitations of currently available imaging techniques with specific emphasis on helical CT. It also presents a novel CADiagnosis scheme for pancreatic tumor segmentation that is based on supervised or unsupervised fuzzy clustering techniques. The proposed algorithm aims at improving pancreatic tumor diagnosis and assessment of treatment effects by automatically segmenting the areas of the pancreas and associated tumor(s) from neighboring organs in CT slices as well as by classifying normal from abnormal pancreatic areas. Preliminary results from a pilot study of the proposed algorithm are presented and discussed including issues of segmentation validation and analysis that are critical to these types of CADiagnosis applications.

Chapter 5 presents the research in the area of flow-mediated dilation (FMD) that offers a mechanism to characterize endothelial function and therefore may play a role in the diagnosis of cardiovascular diseases. Computerized analysis techniques are very desirable to give accuracy and objectivity to the measurements. Virtually all methods proposed up to now to measure FMD rely on accurate edge detection of the arterial wall, and they are not always robust in the

presence of poor image quality or image artifacts. A novel method for automatic dilation assessment based on a global image analysis strategy is presented. We model interframe arterial dilation as a superposition of a rigid motion model and a scaling factor perpendicular to the artery. Rigid motion can be interpreted as a global compensation for patient and probe movements, an aspect that has not been sufficiently studied before. The scaling factor explains arterial dilation. The ultrasound (US) sequence is analyzed in two phases using image registration to recover both transformation models. Temporal continuity in the registration parameters along the sequence is enforced with a Kalman filter since the dilation process is known to be a gradual physiological phenomenon. Comparing automated and gold standard measurements we found a negligible bias (0.04(1.14measurements (bias = 0.47 better reproducibility (CV = 0.46

In Chapter 6 we present the assessment of onset and progression of diseases from images of various modalities is critically dependent on identification of lesions or changes in structures and regions of interest. Mathematical modeling of such discrimination among regions as well as identification of changes in anatomical structures in an image result from the process of segmentation. For clinical applications of segmentation, a compromise between the accuracy and computational speed of segmentation techniques is needed. Optimal segmentation processes based on statistical and adaptive approaches and their applicability to clinical settings have been addressed using diverse modalities of images. Current drawbacks of automated segmentation methodologies stem mostly from nonuniform illumination, inhomogeneous structures, and the presence of noise in acquired images. The effect of preprocessing on the accuracy of segmentation has been discussed. The superior performance of advanced clustering algorithms based on statistical and adaptive approaches over traditional algorithms in medical image segmentation has been presented.

Chapter 7 presents the automatic analysis of color fundus images and with its application to the diagnosis of diabetic retinopathy, a severe and frequent eye disease. We give an overview of computer assistance in this domain and describe in detail some algorithms developed within this framework: the detection of main features in the human eye (vascular tree, the optic disc, and the macula) and the detection of retinal lesions like microaneurysms and hard exudates.

Chapter 8 presents the advanced atherosclerotic plaque that can lead to diseases, such as vessel lumen stenosis, thrombosis, and embolization, which are the leading causes of death and major disability among adults in the United

States. Previous studies have shown that plaque constituents are important determinants for plaque vulnerability and stenosis risk access. To identify and quantitatively measure the composition of atherosclerotic lesions in carotid arteries, plaque segmentation techniques will be discussed in this chapter. First, to extract the lumen contour and outer wall boundary of carotid artery accurately, we will discuss Active Contour Based boundary detection methods, including how to convert exerting energy design and searching process optimizations. Second part is about region-based image segmentation technique, such as Markov random fields, and its applications on image sequence processing. In recent study, plaque components identification with multiple contrast weightings MR images has shown more promising results than single contrast weightings images. In third part, we will introduce multiple contrast weighting MR image segmentation methods and its validation results by comparing with histology images. At last, a software package developed specifically for the quantitative analysis of atherosclerotic plaque by MRI, quantitative vascular analysis system (QVAS) will be presented.

Chapter 9 presents the pre- and postcontrast Gd-DTPA[1] MR images of any body organ hold diagnostic utility in the area of medicine, particularly for breast lesion characterization. This paper reviews the state-of-the-art tools and techniques for lesion characterization, such as uptake curve estimation (functional segmentation), image subtraction, velocity thresholding, differential characteristics of lesions, such as maximum derivative of image sequence, steep slope and washout, fuzzy clustering, Markov random fields, and interactive deformable models such as Live-Wire. In first part of the paper, we discuss the MRI system and breast coils along with the MR breast data acquisition protocol for spatial and temporal MR data collection. Then the perfusion analysis tools are discussed for staging breast tumors. Here, the rate of absorption of the contrast agent (Gad) is used to stage a breast lesion. The differences in contrast enhancement have been shown to be able to help differentiate between benign and malignant lesions. Thus, oncologists, radiologists, and internists have shown great interest in such classification by examining the quantitative characteristics of the tissue signal enhancement. Then we discuss two other major tools for breast lesion characterization. The first set of tools is based on pixel-classification algorithms and second set of tools is based on user-based deformable models such as Live-Wire.

[1] Gadolinium-Dithylene-Triamine-Penta-Acetate, we will refer to it as Gad from now on.

Finally, the paper also presents the user-friendly Marconi Medical System's real-time MR Breast Perfusion[2] Software Analysis System (BPAS), based on Motif using C/C++ and X window libraries that runs on Digital Unix and XP1000 workstations supporting Unix and Linux Operating Systems, respectively. This software was tested on 20 patient studies from the data collected from two major sites in the United States and Europe.

Chapter 10 presents methods for the enhancement and segmentation of 3D local structures, that is, line-like shapes such as blood vessels, sheet-like shapes such as articular cartilage, and blob-like shapes such as nodules in medical volume data. Firstly, a method for enhancement of 3D local structures with various widths is presented. Multiscale Gaussian filters and the eigenvalues of Hessian matrix of the volume function are combined to effectively enhance various widths of structures. The characteristics of multiscale filter responses are analysed to clarify the guidelines for the filter design. Secondly, methods for description and quantification of the 3D local structures are presented. Medial axis/surface elements are locally determined based on the second-order approximations of local intensity structures. Diameter/thickness quantification is performed based on detected medial axis/surface elements. Limits on the accuracy of thickness quantification from 3D MR data is analyzed based on a mathematical models of imaged structures, MR imaging and thickness measurement processes. The utility of the methods is demonstrated by examples using 3D CT and MR data of various parts of the body.

Chapter 11 presents work in the area of CAD design. Research into the computer-aided detection (CAD) of breast lesions from digitised mammograms has been extensive over the past 15 years. The large number of computer algorithms for mammogram contrast enhancement, segmentation, and region discrimination reflects the nontrivial nature in the problem of detecting cancer during breast screening. In addition, due to the enormous variability in the mammographic appearance of the breast, engineering a single CAD solution is formidable. Mammographic CAD must provide a high level of sensitivity for the detection of breast lesions, while maintaining a low number of false-positive regions for each image. This chapter describes an adaptive knowledge-based framework for the detection of breast cancer masses from digitised mammograms. The proposed framework accommodates a set of distinct contrast

[2] Or Breast Uptake.

enhancement and image segmentation experts, used to learn an optimal pipeline of image processing operators for an individual mammogram. It is hypothesised that such an optimal flow will lead to an increase in the sensitivity in the detection of breast lesions. To facilitate efficient training of the adaptive knowledge-based model, a novel method of grouping mammograms on the basis of their mammographic density is proposed. In addition, an automated mechanism for improving the quality of expert radiologist's lesion definitions is presented. To validate this work, 400 digitised mammograms are taken from the publicly available Digital Database of Screening Mammograms (DDSM). The 400 mammograms comprise 200 abnormal and 200 abnormal images complete with lesion ground-truth definitions provided by an expert radiologist. Following the evaluation of the knowledge-based framework in the contrast enhancement and segmentation of mammograms, it is shown that each knowledge-based component out-performs the single best performing expert. Following image segmentation and region prefiltering, a sensitivity of 0.81 with on average 8.65 false-positives per mammogram is reported.

A semiautomatic region growing algorithm, which employs the concept of fuzzy connectedness to perform the simultaneous segmentation of elements of an arbitrary set, is presented in Chapter 12. Because of its general nature, this algorithm can be applied to segmenting dots in the plane, pixels of an image or voxels of a three-dimensional volume. The algorithm is supplied some minimal input by the user and produces an M-segmentation that identifies a grade of membership of every element of the set in each of M objects. The algorithm is illustrated on both mathematically described images and on MR and CT reconstructions.

Chapter 13 presents research in the area of computer-aided diagnosis (CADiagnosis) techniques in medical imaging play the role of a "second opinion" and their goal is to assist the observer in the differenciation between benign and malignant image findings and lesions. CADiagnosis methodologies go beyond the task of automated detection of abnormalities in images and aim at predicting biopsy outcomes from image and/or patient characteristics. In mammography today, CADiagnosis is implemented for masses and calcifications and its output may be a binary one (benign or malignant assignment) or a likelihood (percentage) for a finding to be benign or malignant. This chapter presents a CADiagnosis algorithm developed for assigning a likelihood of malignancy to calcification clusters detected in mammograms. The algorithm is composed of

several modules including image segmentation and classification steps that are based on multiresolution methods and artificial neural networks. Morphology, distribution, and demographics are the domains where features are determined from for the classification task. As a result, the segmentation and feature selection processes of the algorithm are critical in its performance and are the areas we focus on in this chapter. We look particularly into the segmentation aspects of the implementation and the impact of multiresolution filtering has on feature estimation and classification. General aspects of algorithm evaluation and, particularly, segmentation validation are presented using results of experiments conducted for the evaluation of our CADiagnosis scheme as the basis of discussion.

Chapter 14 presents research in the area of neuro segmentation.

Contents

Chapter 1

Model-Based Brain Tissue Classification

Koen Van Leemput,[1] *Dirk Vandermeulen*,[2] *Frederik Maes*,[2]
Siddharth Srivastava,[2] *Emiliano D'Agostino*,[2] *and Paul Suetens*[2]

1.1 Introduction

Several neuropathologies of the central nervous system such as multiple sclerosis (MS), schizophrenia, epilepsy, Alzheimer, and Creutzfeldt–Jakob disease (CJD) are related to morphological and/or functional changes in the brain. Studying such diseases by objectively measuring these changes instead of assessing the clinical symptoms is of great social and economical importance. These changes can be measured in three dimensions in a noninvasive way using current medical imaging modalities. Magnetic resonance imaging (MRI), in particular, is well suited for studying diseases of the nervous system due to its high spatial resolution and the inherent high soft tissue contrast.

Manual analysis of MRI by a trained human expert is a tedious and difficult task, because the structures of interest show complex edge configurations in 3D and may lack clearly visible anatomical borders. In clinical trials, the number of MR images is often so large that manual analysis by human experts is too time-consuming. Furthermore, it is not clear how a human rater combines information obtained from different channels when multispectral MR data are examined. Also, the intra- and interobserver variability associated with manual delineations complicates the analysis of the results. Therefore, there is a need for fully automated methods for MR brain image quantification that can analyze

[1] Department of Radiology, Helsinki University Central Hospital, Finland
[2] Medical Image Computing (Radiology-ESAT/PSI), Faculties of Medicine and Engineering, Katholieke Universiteit Leuven, Belgium

large amounts of multispectral MR data in a reproducible way that correlates well with expert analyses.

A key component in image analysis and interpretation is image segmentation, defined as the delineation of anatomical structures and other regions of interest. In this chapter we will present a framework for the accurate segmentation of brain tissues (Gray matter, white matter, CSF) from multispectral MR images of the brain. We will also discuss how this framework can be used to quantify pathology-related abnormalities (mostly in intensity but also in morphology) within these tissues.

The overall strategy is to build statistical models for normal brain MR images, with emphasis on accurate intensity models. Signal abnormalities are detected as model outliers, i.e., voxels that cannot be well explained by the model. Special attention is paid to automatically estimate all model parameters from the data itself and, hence, to eliminate subjective manual tuning and training.

1.1.1 Segmentation Methodology

The simplest image segmentation methods, such as region growing and edge detection, rely entirely on local image operators and a heuristic grouping of neighboring pixels with similar local photometric characteristics. These approaches are simple to understand and to implement, and are very generic since they do not assume specific knowledge about the objects to be analyzed. However, these methods ultimately fail when either or both of the image data and the object model (shape, context) are complex, as in cross-sectional images of the brain. Indeed, the complex 3-D shape of the brain and its affected areas, and ambiguities in the images induced by the imaging process, such as limited resolution, partial volume effects, noise, low contrast, intensity inhomogeneities, and other artifacts, make brain tissue segmentation difficult, even for human experts.

In order to effectively deal with this complex morphology and MR imaging ambiguity, brain image segmentation methods must incorporate models that describe prior knowledge about the expected geometry and intensity characteristics of the anatomical objects of interest in the images.

- *Intensity-based* methods fit appropriate photometric models to the data. In these approaches, the objects are simple voxels with an associated

scalar, such as the gray value, or vector of characteristic features, such as (ρ, T_1, T_2)-weighted intensity values in MRI. If the feature vectors are represented in a multidimensional feature space, the segmentation strategy then consists of partitioning this feature space in a number of nonoverlapping regions that separate different voxel types. Unclassified voxels then receive the label of their class in feature space. The boundaries between the regions in features space are obtained by optimizing any of a set of decision criteria, depending on the *a priori* assumptions about the feature space distributions. Parametric classification approaches assume that the distributions in feature space follow a certain parametric model. Typically, the multivariate Gaussian model is used. This model can be extended to explicitly include imaging artifacts such as the partial volume effect [1–3] and the intensity inhomogeneity present in MR images [4–6]. Parameter values for the distribution of each class can be learned from a representative set of samples in a supervised training phase, usually requiring cumbersome user interaction. Fully automated, unsupervised learning procedures estimate the distribution parameters from the image to be segmented itself.

- *Geometry-based* methods use prior knowledge about the overall shape of an object to separate it from its surroundings in the image. Typically, a surface deforms under the influence of external image derived forces (attracting the surface to object edges, etc.) and internal elasticity constraints (e.g. surface continuity and smoothness) [7]. An extensive survey of these methods in medical image analysis is given in [8]; recent examples include [9–12]. Within the context of brain image analysis, deformable brain atlas-guided approaches have been proposed. Here, prior knowledge about the image scene is represented iconically (i.e. an image-like representation). The image is segmented by matching it with the iconic atlas representation. The matching process must have sufficient degrees of freedom and robustness so as to cope with the biological variability and pathological abnormalities. However, even with nonlinear registration methods, accurate brain tissue segmentations are difficult due to anatomical variability of the cortical folding.

Intensity-based tissue classification and geometry-based atlas-driven methods are seemingly complementary segmentation strategies.

- The advantage of MR-intensity-based tissue classification is its ability to produce high-quality definitions of tissue boundaries. This is especially important for human brain tissue classification, where highly curved interfaces between tissues (such as between gray and white matter) are difficult to recover from finite resolution images. However, it is unsuccessful when different structures have overlapping feature distributions (e.g., brain tissue and extracranial tissue in T1-weighted MRI).

- Geometry-based methods have been successfully applied to the localization of particular anatomical structures, where sufficient information on shape and context is available. These methods, however, often require accurate initialization and, more importantly, can fail in the case of highly variable structures such as the cerebral cortex and in the presence of abnormal anatomical structures.

Following Warfield *et al.* [13] and Collins *et al.* [14], who developed the idea that the inherent limitations of intensity-based classification can be alleviated by combining it with an elastically deformable atlas, we propose here to combine photometric and geometric models in a single framework. Tissues are classified using an unsupervised parametric approach by using a mixture of Gaussians as the feature space distribution model in an iterative expectation-maximization loop. Classifications are further initialized and constrained by iconic matching of the image to a digital atlas containing spatial maps of prior tissue probabilities.

Section 1.1.2 presents the standard intensity model and optimization approach that we will use throughout this chapter. Section 1.1.3 discusses the basic geometric model (a brain atlas) and its iconic matching to image data using linear and nonlinear registration algorithms. Section 1.1.4 summarizes the changes made to the basic intensity model in order to model the MR imaging process more faithfully and to make the segmentation procedure more robust in the presence of abnormalities.

1.1.2 Intensity Model and the Expectation-Maximization (EM) Algorithm

The intensity model proposed here is the so-called mixture of normal distributions [15–17]. Let $\boldsymbol{Y} = \{y_j, j = 1, 2, \ldots, J\}$ be a C-channel MR image with a

total of J voxels, where $\boldsymbol{y}_j$ denotes the possibly multispectral intensity of voxel j. Suppose that there are K tissue types present in the imaged area, and let $l_j \in \{1, 2, \ldots, K\}$ denote the tissue type to which voxel j belongs. In the mixture model, it is assumed that each tissue type $\boldsymbol{k}$ has a typical intensity $\boldsymbol{\mu}_k$ in the image, with tissue-specific normally distributed intensity fluctuations in the voxels. In other words, the probability density that voxel j of tissue type l_j has intensity $\boldsymbol{y}_j$ is given by

$$f(\boldsymbol{y}_j \mid l_j, \Phi) = G_{\boldsymbol{\Sigma}_{l_j}}(\boldsymbol{y}_j - \boldsymbol{\mu}_{l_j}) \tag{1.1}$$

Here $G_{\boldsymbol{\Sigma}}(\cdot)$ denotes a zero-mean normal distribution with covariance $\boldsymbol{\Sigma}$, and $\Phi = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \boldsymbol{k} = 1, 2, \ldots, K\}$ represents the total set of model parameters.

For notational convenience, let all the voxel labels l_j be grouped in a label image $\boldsymbol{L} = \{l_j, j = 1, 2, \ldots, J\}$. It is assumed that the label l_j of each voxel is drawn independently from the labels of the other voxels, with an *a priori* known probability π_k, i.e.

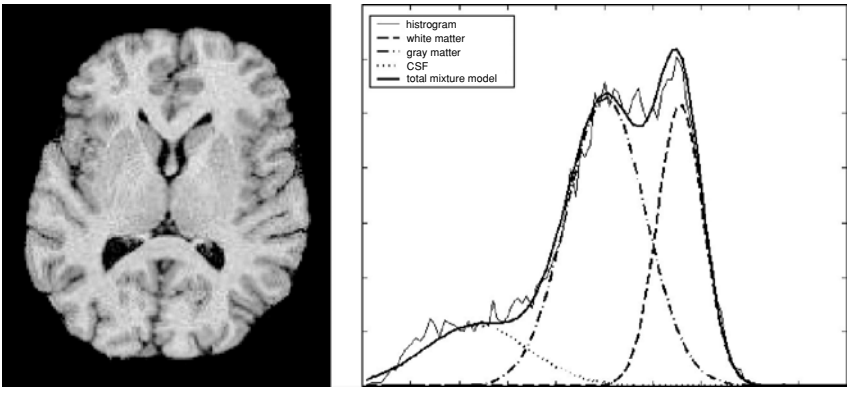
$$f(\boldsymbol{L}) = \prod_j \pi_{l_j} \tag{1.2}$$

The overall probability density for image $\boldsymbol{Y}$ given the model parameters Φ is then given by

$$\begin{aligned} f(\boldsymbol{Y} \mid \Phi) &= \sum_{\boldsymbol{L}} f(\boldsymbol{Y} \mid \boldsymbol{L}, \Phi) f(\boldsymbol{L}) \\ &= \prod_{\boldsymbol{j}} f(\boldsymbol{y_j} | \Phi) \\ \text{with } f(\boldsymbol{y_j} \mid \Phi) &= \sum_{\boldsymbol{k}} f(\boldsymbol{y_j} \mid \boldsymbol{l_j} = \boldsymbol{k}, \Phi) \cdot \boldsymbol{\pi_k} \end{aligned} \tag{1.3}$$

Equation (1.3) is the well-known mixture model (see Fig. 1.1). It models the histogram of image intensities as a sum of normal distributions, each distribution weighted with its prior probability $\boldsymbol{\pi_k}$.

Image segmentation aims to reconstruct the underlying tissue labels $\boldsymbol{L}$ based on the image $\boldsymbol{Y}$. If an estimation of the model parameters is somehow available, then each voxel can be assigned to the tissue type that best explains its intensity. Unfortunately, the result depends largely on the model parameters used. Typically, these are estimated by manually selecting representative points in the image of each of the classes considered. However, once all the voxels are classified, the model parameter estimation can in its turn automatically be improved based on all the voxels instead of on the subjectively selected ones alone.



(a) (b)

Figure 1.1: The mixture model fitted to a T1-weighted brain MR image: (a) the intracranial volume and (b) its intensity histogram with a mixture of normal distributions overlayed. The normal distributions correspond to white matter, gray matter, and CSF.

Intuitively, both the segmentation and the model parameters can be estimated in a more objective way by interleaving the segmentation with the estimation of the model parameters.

The expectation-maximization (EM) algorithm [18] formalizes this intuitive approach. It estimates the maximum likelihood (ML) parameters $\hat{\Phi}$

$$\hat{\Phi} = \arg\max_{\Phi} \log f(\boldsymbol{Y} \mid \Phi)$$

by iteratively filling in the unknown tissue labels $\boldsymbol{L}$ based on the current parameter estimation Φ, and recalculating Φ that maximizes the likelihood of the so-called complete data $\{\boldsymbol{Y}, \boldsymbol{L}\}$. More specifically, the algorithm interleaves two steps:

Expectation step: find the function

$$Q(\Phi \mid \Phi^{(m-1)}) = E_{\boldsymbol{L}}[\log f(\boldsymbol{Y}, \boldsymbol{L} \mid \Phi) \mid \boldsymbol{Y}, \Phi^{(m-1)}]$$

Maximization step: find

$$\Phi^{(m)} = \arg\max_{\Phi} Q(\Phi \mid \Phi^{(m-1)})$$

with m the iteration number. It has been shown that the likelihood $\log f(\boldsymbol{Y} \mid \Phi)$ is guaranteed to increase at each iteration for EM algorithms [19].

With the image model described above, the expectation step results in a statistical classification of the image voxels

$$f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)}) = \frac{f(\boldsymbol{y_j} \mid l_j, \Phi^{(m-1)}) \cdot \pi_{l_j}}{\sum_k f(\boldsymbol{y_j} \mid l_j = k, \Phi^{(m-1)}) \cdot \pi_k} \tag{1.4}$$

and the subsequent maximization step involves

$$\boldsymbol{\mu_k}^{(m)} = \frac{\sum_j f(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)}) \cdot y_j^{(m-1)}}{\sum_j f(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)})} \tag{1.5}$$

$$\Sigma_k^{(m)} = \frac{\sum_j f(l_j = \boldsymbol{k} \mid \boldsymbol{Y}, \Phi^{(m-1)}) \cdot (\boldsymbol{y_j}^{(m-1)} - \boldsymbol{\mu_k}^{(m)}) \cdot (\boldsymbol{y_j}^{(m-1)} - \boldsymbol{\mu_k}^{(m)})^t}{\sum_j f(l_j = \boldsymbol{k} \mid \boldsymbol{Y}, \Phi^{(m-1)})} \tag{1.6}$$

Thus, the algorithm iteratively improves the model parameters by interleaving two steps (see Fig. 1.2): classification of the voxels based on the estimation of the normal distributions (Eq. 1.4) and estimation of the normal distributions based on the classification (Eqs. 1.5 and 1.6). Upon convergence, Eq. (1.4) yields the final classification result.

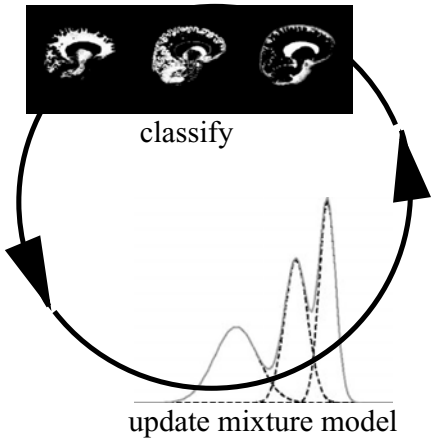
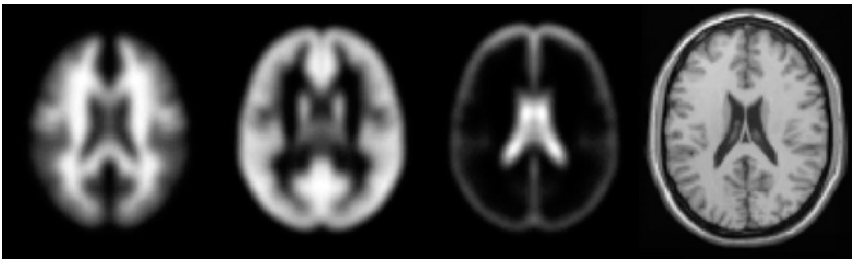


Figure 1.2: Estimating the model parameters of the mixture model with an expectation-maximization algorithm results in an iterative two-step process that interleaves classification of the voxels with reestimation of the normal distributions.

1.1.3 Geometry Model and the MMI Algorithm

The iterative model parameter estimation algorithm described above needs to be initialized with a first estimate of the parameters. One possibility to obtain such an estimate is to have the user manually select voxels that are representative for each of the classes considered. However, to eliminate the variability induced by such a preceding training phase, we avoid manual intervention by the use of a digital brain atlas that contains spatially varying prior probability maps for the location of white matter, gray matter, and CSF (see Fig. 1.3). These probability maps were obtained by averaging binary white matter, gray matter, and CSF segmentations of MR brain images from a large number of subjects, after normalization of all images into the same space using an affine transformation [20]. To apply this *a priori* information, the atlas is first normalized to the space of the study image by matching the study image to a T1 template associated with the atlas (see Fig. 1.3) by using the affine multimodality registration technique based on maximization of mutual information (MMI) of Maes *et al.* [21].

Mutual information (MI) is a basic concept from information theory, which is applied in the context of image registration to measure the amount of information that one image contains about the other. The MMI registration criterion postulates that MI is maximal when the images are correctly aligned. Mutual information does not rely on the intensity values directly to measure correspondence between different images, but on their relative occurrence in each of the



(a) (b) (c) (d)

Figure 1.3: Digital brain atlas with spatially varying *a priori* probability maps for (a) white matter, (b) gray matter, and (c) CSF. High intensities indicate high *a priori* probabilities. The atlas also contains a T1 template image (d), which is used for registration of the study images to the space of the atlas. (Source: Ref. [23].)

images separately and co-occurrence in both images combined. Unlike other voxel-based registration criteria, based on for instance intensity differences or intensity correlation, the MI criterion does not make limiting assumptions about the nature of the relationship between the image intensities of corresponding voxels in the different modalities, which is highly data-dependent, and does not impose constraints on the image content of the modalities involved. This explains the success of MMI for multimodal image registration in a wide range of applications involving various modality combinations [22]. It has furthermore been shown [21] that this registration criterion is fairly insensitive to moderate bias fields, such that it can be applied fully automatically and reliably to the MR images with a bias field inhomogeneity (see section 1.2). The properly registered and reformatted *a priori* tissue probability maps of the atlas provide an initial estimate of the classification from which initial values for the class-specific distribution parameters $\boldsymbol{\mu_k}$ and $\boldsymbol{\Sigma_k}$ can be computed. This approach frees us from having to interactively indicate representative voxels of each class, which makes our method more objective and reproducible and allows the method to be fully automated.

The classification and intensity distribution parameters are then updated using the iterative scheme based on the EM procedure as outlined above. During iterations, the atlas is further used to spatially constrain the classification by assigning its prior probability maps to the *a priori* class probabilities π_k. Thus, the voxels are not only classified based on their intensities, but also based on their spatial position. This makes the algorithm more robust, especially when the images are corrupted with a heavy bias field.

However, in the presence of gross morphological differences between atlas and patient images, intensity-based atlas-guided pixel classification as described above, using only affine iconic matching of atlas to image, may fail. Figure 1.4 (left) shows a cross section through the brain of a patient affected by periventricular leukomalacia (PVL). This brain presents gross morphological differences compared to a normal brain, especially around the ventricles, which are strongly enlarged. Brain tissue segmentation of such images using affine iconic matching of atlas and patient images will fail, if the gross morphological differences between atlas and patient images are not corrected for. Indeed, in this particular case, the affinely matched atlas labels large portions of the enlarged ventricles as white matter. The initial estimate of the tissue intensity parameters (i.e. mean and variance) is thus not reliable and it is therefore unlikely that the iterative

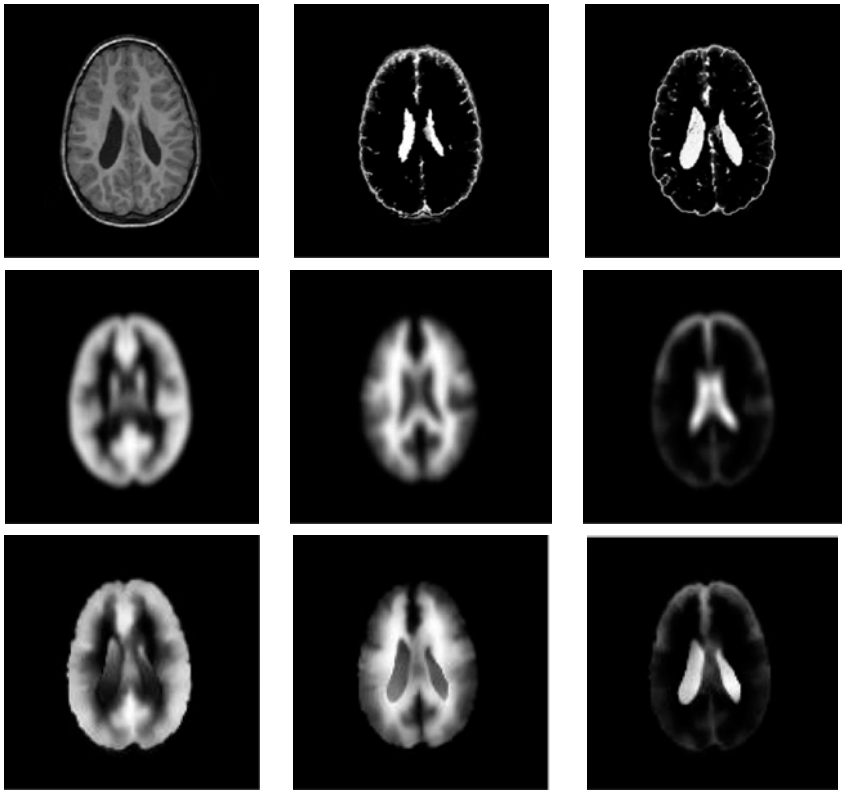


Figure 1.4: Top row (from left to right): Original T1 MPRAGE patient image; CSF segmented by atlas-guided intensity-based tissue classification using affine registration to atlas template; CSF segmented after nonrigid matching of the atlas. Middle row: Atlas priors for gray matter, white matter, and CSF affinely matched to patient image. Bottom row: *Idem* after nonrigid matching.

segmentation process, which alternates between pixel classification and parameter estimation, converges to the correct segmentation solution (Fig. 1.4 (top row, middle)). Nonrigid registration of atlas and patient images using the method presented in [24] can better cope with this extra variability. Note how the segmentation of the enlarged ventricles is much improved (see Fig. 1.4 (top row, right)).

1.1.4 Model-Based Brain Tissue Classification: Overview

The standard intensity model described in section 1.1.2 can be extended in several ways in order to better represent real MR images of the brain. The overall

strategy is to build increasingly better statistical models for normal brain MR images and to detect disease-related (e.g. MS or CJD) signal abnormalities as voxels that cannot be well explained by the model. Figure 1.5 shows typical samples of the proposed models in increasing order of complexity, starting from the original mixture model, shown in Fig. 1.5(a). For each of the models, the same EM framework is applied to estimate the model parameters. An unusual aspect of the methods presented here is that all model parameters are estimated from the data itself, starting from an initialization that is obtained without user intervention. This ensures that the models retrain themselves fully automatically on each individual scan, allowing the method to analyze images with previously unseen MR pulse sequences and voxel sizes. In this way, subjective manual tuning and training is eliminated, which would make the results not fully reproducible.

The first problem for the segmentation technique of section 1.1.2 is the corruption of MR images with a smoothly varying intensity inhomogeneity or bias field [25, 26], which results in a nonuniform intensity of the tissues over the image area as shown in Fig. 1.6. This bias is inherent to MR imaging and is caused by equipment limitations and patient-induced electrodynamic interactions [26]. Although not always visible for a human observer, it can cause serious misclassifications when intensity-based segmentation techniques are used. In section 1.2 the mixture model is therefore extended by explicitly including a parametric model for the bias field. Figure 1.5(b) shows a typical sample of the resulting model. The model parameters are then iteratively estimated by interleaving three steps: classification of the voxels; estimation of the normal distributions; and estimation of the bias field. The algorithm is initialized with information from a digital brain atlas about the *a priori* expected location of tissue classes. This allows full automation of the method without need for user interaction, yielding fully objective and reproducible results.

As a consequence of the assumption that the tissue type of each voxel is independent from the tissue type of the other voxels, each voxel is classified independently, based only on its intensity. However, the intensity of some tissues surrounding the brain is often very similar to that of brain tissue, which makes a correct classification based on intensity alone impossible. Therefore, the model is further extended in section 1.3 by introducing a Markov random field (MRF) prior on the underlying tissue labels of the voxels. Such a MRF takes into account that the various tissues in brain MR images are not just randomly distributed over the image area, but occur in clustered regions of homogeneous tissue. This

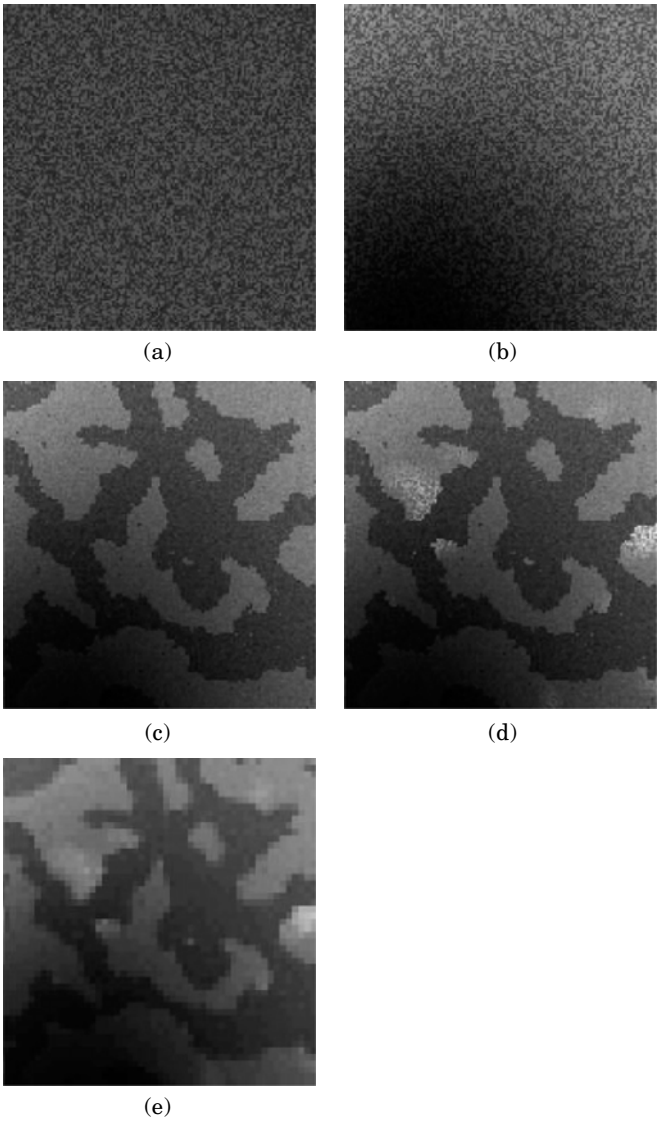
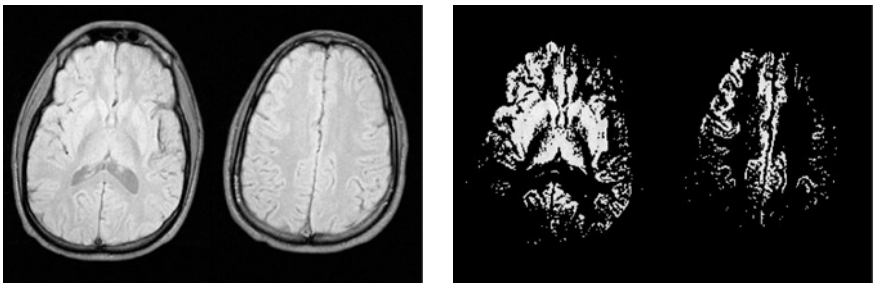


Figure 1.5: Illustration of the statistical models for brain MR images used in this study. The mixture model (a) is first extended with an explicit parametric model for the MR bias field (b). Subsequently, an improved spatial model is used that takes into account that tissues occur in clustered regions in the images (c). Then the presence of pathological tissues, which are not included in the statistical model, is considered (d). Finally, a downsampling step introduces partial voluming in the model (e).

is illustrated in Fig. 1.5(c), which shows a sample of the total resulting image model. The MRF brings general spatial and anatomical constraints into account during the classification, facilitating discrimination between tissue types with similar intensities such as brain and nonbrain tissues.

The method is further extended in section 1.4 in order to quantify lesions or disease-related signal abnormalities in the images (see Fig. 1.5(d)). Adding an explicit model for the pathological tissues is difficult because of the wide variety of their appearance in MR images, and because not every individual scan contains sufficient pathology for estimating the model parameters. These problems are circumvented by detecting lesions as voxels that are not well explained by the statistical model for normal brain MR images. Based on principles borrowed from the robust statistics literature, tissue-specific voxel weights are introduced that reflect the typicality of the voxels in each tissue type. Inclusion of these weights results in a robustized algorithm that simultaneously detects lesions as model outliers and excludes these outliers from the model parameter estimation. In section 1.5, this outlier detection scheme is applied for fully automatic segmentation of MS lesions from brain MR scans. The method is validated by comparing the automatic lesions segmentations to manual tracings by human experts.

Thus far, the intensity model assumes that each voxel belongs to one single tissue type. Because of the complex shape of the brain and the finite resolution of the images, a large part of the voxels lies on the border between two or more tissue types. Such border voxels are commonly referred to as partial volume (PV) voxels as they contain a mixture of several tissues at the same time. In order to be able to accurately segment major tissue classes as well as to detect the subtle signal abnormalities in MS, e.g., the model for normal brain MR images can be further refined by explicitly taking this PV effect into account. This is accomplished by introducing a downsampling step in the image model, adding up the contribution of a number of underlying subvoxels to form the intensity of a voxel. In voxels where not all subvoxels belong to the same tissue type, this causes partial voluming, as can be seen in Fig. 1.5(e). The derived EM algorithm for estimating the model parameters provides a general framework for partial volume segmentation that encompasses and extends existing techniques. A full presentation of this PV model is outside the scope of this chapter, however. The reader is referred to [27] for further details.



(a) (b)

Figure 1.6: The MR bias field in a proton density-weighted image. (a) Two axial slices; (b) the same slices after intensity thresholding.

1.2 Automated Bias Field Correction

A major problem for automated MR image segmentation is the corruption with a smoothly varying intensity inhomogeneity or bias field [25, 26]. This bias is inherent to MR imaging and is caused by equipment limitations and patient-induced electrodynamic interactions. Although not always visible for a human observer, as illustrated in Fig. 1.6, correcting the image intensities for bias field inhomogeneity is a necessary requirement for robust intensity-based image analysis techniques. Early methods for bias field estimation and correction used phantoms to empirically measure the bias field inhomogeneity [28]. However, this approach assumes that the bias field is patient independent, which it is not [26]. Furthermore, it is required that the phantom's scan parameters are the same as the patient's, making this technique impractical and even useless as a retrospective bias correction method. In a similar vein, bias correction methods have been proposed for surface coil MR imaging using an analytic correction of the MR antenna reception profile [29], but these suffer from the same drawbacks as do the phantom-based methods. Another approach, using homomorphic filtering [30], assumes that the frequency spectrum of the bias field and the image structures are well separated, but this assumption is generally not valid for MR images [28, 31].

While bias field correction is needed for good segmentation, many approaches have exploited the idea that a good segmentation helps to estimate the bias field. Dawant *et al.* [31] manually selected some points inside white matter and estimated the bias field as the least-squares spline fit to the

intensities of these points. They also presented a slightly different version where the reference points are obtained by an intermediate classification operation, using the estimated bias field for final classification. Meyer *et al.* [32] also estimated the bias field from an intermediate segmentation, but they allowed a region of the same tissue type to be broken up into several subregions which creates additional but sometimes undesired degrees of freedom.

Wells *et al.* [4] described an iterative method that interleaves classification with bias field correction based on ML parameter estimation using an EM algorithm. However, for each set of similar scans to be processed, their method, as well as its refinement by other authors [5, 6], needs to be supplied with specific tissue class conditional intensity models. Such models are typically constructed by manually selecting representative points of each of the classes considered, which may result in segmentations that are not fully objective and reproducible.

In contrast, the method presented here (and in [23]) does not require such a preceding training phase. Instead, a digital brain atlas is used with *a priori* probability maps for each tissue class to automatically construct intensity models for each individual scan being processed. This results in a fully automated algorithm that interleaves classification, bias field estimation, and estimation of class-conditional intensity distribution parameters.

1.2.1 Image Model and Parameter Estimation

The mixture model outlined in section 1.1.2 is extended to include a model for the bias field. We model the spatially smoothly varying bias fields as a linear combination of P polynomial basis functions $\phi_p(\boldsymbol{x}_j)$, $p = 1, 2, \ldots, P$, where $\boldsymbol{x}_j$ denotes the spatial position of voxel j. Not the observed intensities $\boldsymbol{y}_j$ but the bias corrected intensities $\boldsymbol{u}_j$ are now assumed to be distributed according to a mixture of class-specific normal distributions, such that Eq. (1.1) above is replaced by

$$f(\boldsymbol{y}_j \mid l_j, \Phi) = G_{\Sigma_{l_j}}(\boldsymbol{u}_j - \boldsymbol{\mu}_{l_j})$$

$$\boldsymbol{u}_j = \boldsymbol{y}_j - [\boldsymbol{b}_1 \cdots \boldsymbol{b}_C]^t \begin{bmatrix} \phi_1(\boldsymbol{x}_j) \\ \vdots \\ \phi_P(\boldsymbol{x}_j) \end{bmatrix}$$

with $\boldsymbol{b}_c, c = 1, 2, \ldots, C$ indicating the bias field parameters of MR channel c, and $\Phi = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \boldsymbol{b}_c, \boldsymbol{k} = 1, 2, \ldots, K, c = 1, 2, \ldots, C\}$ as the total set of model parameters.

With the addition of the bias field model, estimation of the model parameters with an EM algorithm results in an iterative procedure that now interleaves three steps (see Fig. 1.7): classification of the image voxels (Eq. 1.4); estimation of the normal distributions (Eqs. (1.5) and (1.6) but with the bias corrected intensities $\boldsymbol{u}_j$ replacing the original intensities $\boldsymbol{y}_j$); and estimation of the bias field. For the unispectral case, the bias field parameters are given by the following expression[3]:

$$\boldsymbol{b}^{(m)} = (\boldsymbol{A}^t \boldsymbol{W}^{(m)} \boldsymbol{A})^{-1} \boldsymbol{A}^t \boldsymbol{W}^{(m)} \boldsymbol{R}^{(m)}$$

with

$$\boldsymbol{A} = \begin{bmatrix} \phi_1(\boldsymbol{x}_1) & \phi_2(\boldsymbol{x}_1) & \phi_3(\boldsymbol{x}_1) & \cdots \\ \phi_1(\boldsymbol{x}_2) & \phi_2(\boldsymbol{x}_2) & \phi_3(\boldsymbol{x}_2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$\boldsymbol{W}^{(m)} = \operatorname{diag}\big(w_j^{(m)}\big), \qquad w_j^{(m)} = \sum_k w_{jk}^{(m)}, \qquad w_{jk}^{(m)} = \frac{f\big(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)}\big)}{\sigma_k^2}$$

$$\boldsymbol{R}^{(m)} = \begin{bmatrix} y_1 - \tilde{y}_1^{(m)} \\ y_2 - \tilde{y}_2^{(m)} \\ \vdots \end{bmatrix}, \qquad \tilde{y}_j^{(m)} = \frac{\sum_k w_{jk}^{(m)} \mu_k^{(m)}}{\sum_k w_{jk}^{(m)}}$$

This can be interpreted as follows (see Fig. 1.8). Based on the current classification and distribution estimation, a prediction $\{\tilde{y}_j, j = 1, 2, \ldots, J\}$ of the MR image without the bias field is constructed (Fig. 1.8(b)). A residue image $\boldsymbol{R}$ (Fig. 1.8(c)) is obtained by subtracting this predicted signal from the original image (Fig. 1.8(a)). The bias (Fig. 1.8(e)) is then estimated as a weighted least-squares fit through the residue image using the weights $\boldsymbol{W}$ (Fig. 1.8(d)), each voxel's weight being inversely proportional to the variance of the class that voxel belongs to. As can be seen from Fig. 1.8(d), the bias field is therefore computed primarily from voxels that belong to classes with a narrow intensity distribution, such as white and gray matter. The smooth spatial model extrapolates the bias

[3] For more general expressions for the multispectral case we refer to [23]; in the unispectral case $\boldsymbol{y}_j = y_j$, $\boldsymbol{\mu}_k = \mu_k$, and $\boldsymbol{\Sigma}_k = \sigma_k^2$.

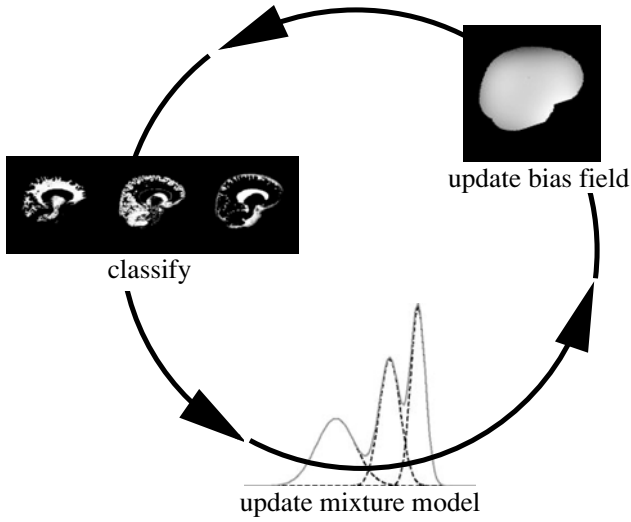


Figure 1.7: Adding a model for the bias field results in a three-step expectation-maximization algorithm that iteratively interleaves classification, estimation of the normal distributions, and bias field correction.

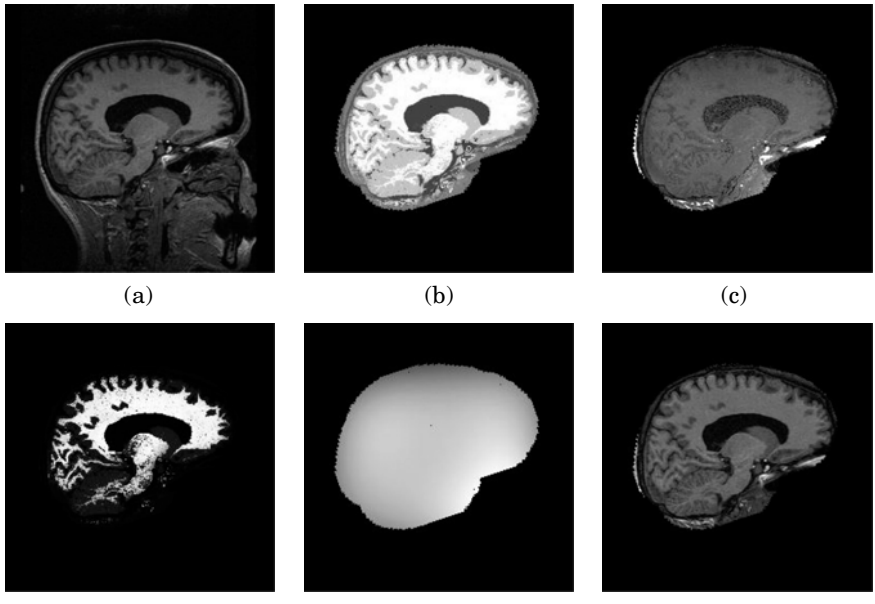


Figure 1.8: Illustration of the bias correction step on a 2D slice of a T1-weighted MR image: (a) original image; (b) predicted signal based on previous iterations; (c) residue image; (d) weights; (e) estimated bias field; (f) corrected image. (Source: Ref. [23].)

field from these regions, where it can be confidently estimated from the data, to regions where such estimate is ill-conditioned (CSF, nonbrain tissues).

1.2.2 Examples and Discussion

Examples of the performance of the method are shown in Figs. 1.9 and 1.10. Figure 1.9 depicts the classification of a high-resolution sagittal T1-weighted MR image, both for the original two-step algorithm without bias correction of section 1.1.2 and for the new three-step algorithm with bias correction. Because a relatively strong bias field reduces the intensities at the top of the head, bias correction is necessary as white matter is wrongly classified as gray matter in that area otherwise. Figure 1.10 clearly shows the efficiency of the bias correction on a 2-D multislice T1-weighted image. Such multislice images are acquired in an interleaved way, and are typically corrupted with a slice-by-slice constant intensity offset, commonly attributed to gradient eddy currents and crosstalk

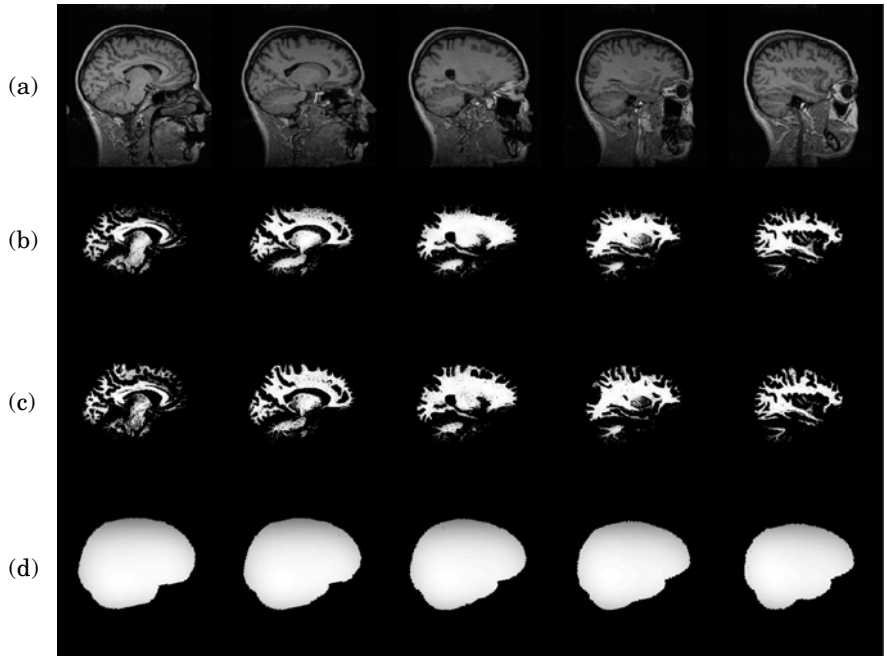


Figure 1.9: Slices of a high-resolution T1-weighted MR image illustrating the performance of the method: (a) original data; (b) white matter classification without bias correction; (c) white matter classification with bias correction; (d) estimated bias field. (Source: Ref. [23].)

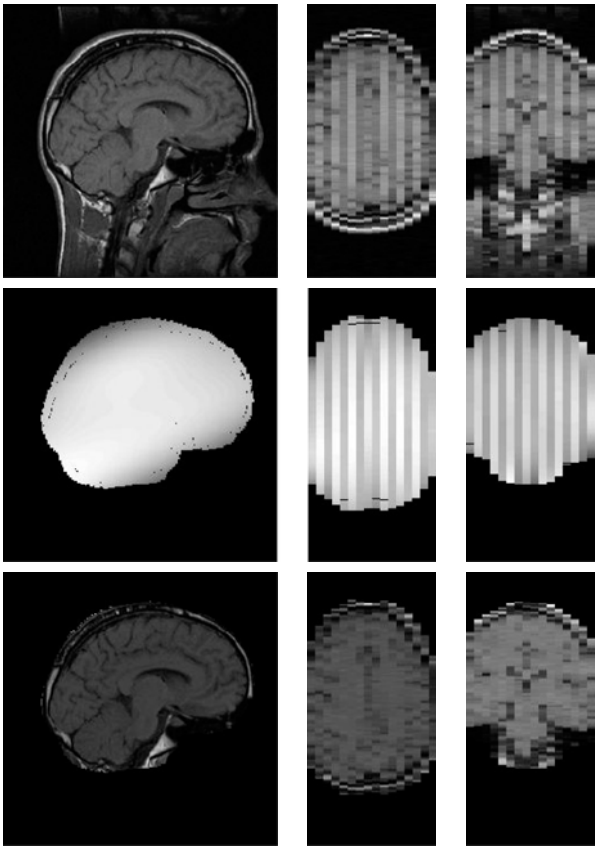


Figure 1.10: An example of bias correction of a T1-weighted 2-D multislice image corrupted with slice-by-slice offsets. From top row to bottom row: original data, estimated bias, corrected data. (Source: Ref. [23].)

between slices [25], and clearly visible as an interleaved bright-dark intensity pattern in a cross-section orthogonal to the slices in Fig. 1.10.

In earlier EM approaches for bias correction, the class-specific intensity distribution parameters $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ were determined by manual training and kept fixed during the iterations [4–6]. It has been reported [6, 33, 34] that these methods are sensitive to the training of the classifier, i.e., they produce different results depending on which voxels were selected for training. In contrast, our algorithm estimates its tissue-specific intensity distributions fully automatically on each individual scan being processed, starting from a digital brain atlas. This avoids all manual intervention, yielding fully objective and reproducible results. Moreover, it eliminates the danger of overlooking some tissue types during a manual training phase, which is typically a problem in regions surrounding the

brain, consisting of several different tissues, and which may cause severe errors in the residue image and the bias field estimation [6, 34]. Guillemaud and Brady [6] proposed to model nonbrain tissues by a single class with a uniform distribution, artificially assigning the nonbrain tissue voxels a zero weight for the bias estimation. This is not necessary with our algorithm: the class distribution parameters are updated at each iteration from all voxels in the image and classes consisting of different tissue types are automatically assigned a large variance. Since the voxel weights for the bias correction are inversely proportional to the variance of the class each voxel is classified to, such tissues are therefore automatically assigned a low weight for the bias estimation.

A number of authors have proposed bias correction methods that do not use an intermediate classification. Styner, Brechbühler *et al.* [33, 35] find the bias field for which as many voxels as possible have an intensity in the corrected image that lies close to that of a number of predefined tissue types. Other approaches search for the bias field that makes the histogram of the corrected image as sharp as possible [36–38]. The method of Sled *et al.* [36] for instance is based on deconvolution of the histogram of the measured signal, assuming that the histogram of the bias field is Gaussian, while Mangin [37] and Likar *et al.* [38] use entropy to measure the histogram sharpness. Contrary to our approach, these methods treat all voxels alike for bias estimation. This looks rather unnatural, since it is obvious that the white matter voxels, which have a narrow intensity histogram, are much more suited for bias estimation than, for instance, the tissues surrounding the brain or ventricular CSF. As argued above, our algorithm takes this explicitly into account by the class-dependent weights assigned to each voxel. Furthermore, lesions can be so large in a scan of a MS patient that the histogram of the corrected image may be sharpest when the estimated bias field follows the anatomical distribution of the lesions. As will be shown in section 1.4, our method can be made robust for the presence of such pathologic tissues in the images, estimating the bias field only from normal brain tissue.

1.3 Modeling Spatial Context

As a consequence of the assumption that the tissue type of each voxel is independent from the tissue type of the other voxels, each voxel is classified independently based on its intensity. This yields acceptable classification results as

long as the different classes are well separated in intensity feature space, i.e. have a clearly discernible associated intensity distribution. Unfortunately, this is not always true for MR images of the brain, especially not when only one MR channel is available. While white matter, gray matter, and CSF usually have a characteristic intensity, voxels surrounding the brain often show an MR intensity that is very similar to brain tissue. This may result in erroneous classifications of small regions surrounding the brain as gray matter or white matter.

We therefore extend the model with a Markov random field prior (MRF), introducing general spatial and anatomical constraints during the classification [39]. The MRF is designed to facilitate discrimination between brain and nonbrain tissues while preserving the detailed interfaces between the various tissue classes within the brain.

1.3.1 Regularization Using a MRF Model

Previously, it was assumed that the label l_j of each voxel is drawn independently from the labels of the other voxels, leading to Eq. (1.2) for the prior probability distribution for the underlying label image $\boldsymbol{L}$. Now a more complex model for $\boldsymbol{L}$ is used, more specifically a MRF. Such a MRF model assumes that the probability that voxel $\boldsymbol{j}$ belongs to tissue type k depends on the tissue type of its neighbors. The Hammersley–Clifford theorem states that such a random field is a Gibbs random field (see [40] and the references therein), i.e. its configurations obey a Gibbs distribution

$$f(\boldsymbol{L} \mid \Phi) = Z(\Phi)^{-1} \exp[-U(\boldsymbol{L} \mid \Phi)]$$

where $Z(\Phi) = \boldsymbol{\Sigma}_{\boldsymbol{L}} \exp[-U(\boldsymbol{L} \mid \Phi)]$ is a normalization constant called the partition function and $U(\boldsymbol{L} \mid \Phi)$ is an energy function dependent on the model parameters Φ.

A simple MRF is used that is defined on a so-called first-order neighborhood system, i.e. only the six nearest neighbors on the 3-D-image lattice are used. Let $\mathcal{N}_j^p$ denote the set of the four neighbors of voxel $\boldsymbol{j}$ in the plane and $\mathcal{N}_j^o$ its two neighbors out of the plane. Since the voxel size in the z direction is usually different from the within-plane voxel size in MR images, the following Potts model (the extension of the binary Ising model [41] to more than two classes)

is used:

$$U(\boldsymbol{L} \mid \Phi) = \frac{1}{2} \sum_j \left(\sum_{j' \in \mathcal{N}_j^p} \xi_{l_j l_{j'}} + \sum_{j' \in \mathcal{N}_j^o} \nu_{l_j l_{j'}} \right)$$

Here, the MRF parameters $\xi_{kk'}$ and $\nu_{kk'}$ denote the cost associated with transition from class k to class k' among neighboring voxels in the plane and out of the plane, respectively. If these costs are higher for neighbors belonging to different classes than for neighbors of the same tissue class, the MRF favors configurations of $\boldsymbol{L}$ where each tissue is spatially clustered. An example of this is shown in Fig. 1.5(c).

It has been described in the literature that fine structures, such as the interface between white matter and gray matter in the brain MR images, can be erased by the Potts/Ising MRF model [5, 42]. The MRF may overregularize such subtle borders and tempt to produce nicely smooth interfaces. Therefore, a modification is used that penalizes anatomically impossible combinations such as a gray matter voxel surrounded by nonbrain tissues, while at the same time preserving edges between tissues that are known to border each other. We impose that a voxel surrounded by white matter and gray matter voxels must have the same *a priori* probability for white matter as for gray matter by adding appropriate constraints on the MRF transition costs $\xi_{kk'}$ and $\nu_{kk'}$. As a result, voxels surrounded by brain tissues have a low probability for CSF and other nonbrain tissues, and a high but equal probability for white and gray matter. The actual decision between white and gray matter is therefore based only on the intensity, so that the interface between white and gray matter is unaffected by the MRF. Similar constraints are applied for other interfaces as well.

Since the voxel labels are not independent with this model, the expectation step of the EM algorithm no longer yields the classification of equation 4. Because of the interaction between the voxels, the exact calculation of $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)})$ involves calculation of all the possible realizations of the MRF, which is not computationally feasible. Therefore, an approximation was adopted that was proposed by Zhang [43] and Langan *et al.* [44], based on the mean field theory from statistical mechanics. This mean field approach suggests an approximation to $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)})$ based on the assumption that the influence of $l_{j'}$ of all other voxels $j' \neq j$ in the calculation of $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)})$ can be approximated by the influence of their classification $f(l_{j'} \mid \boldsymbol{Y}, \Phi^{(m-2)})$ from the previous iteration.

This yields

$$f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)}) \approx \frac{f(\boldsymbol{y}_j \mid l_j, \Phi^{(m-1)}) \cdot \pi_j^{(m-1)}(l_j)}{\Sigma_k f(\boldsymbol{y}_j \mid l_j = \boldsymbol{k}, \Phi^{(m-1)}) \cdot \pi_j^{(m-1)}(k)} \tag{1.7}$$

for the classification, where

$$\pi_j^{(m-1)}(l_j) \sim \exp\left(- \sum_{j' \in \mathcal{N}_j^p} \sum_k f(l_{j'} = \boldsymbol{k} \mid \boldsymbol{Y}, \Phi^{(m-2)}) \cdot \xi_{l_j k} - \sum_{j' \in \mathcal{N}_j^o} \sum_k f(l_{j'} = \boldsymbol{k} \mid \boldsymbol{Y}, \Phi^{(m-2)}) \cdot \nu_{l_{j_k}} \right)$$

The difference with Eq. (1.4) lies herein that now the *a priori* probability that a voxel belongs to a specific tissue class depends on the classification of its neighboring voxels.

With the addition of the MRF, the subsequent maximization step in the EM algorithm not only involves updating the intensity distributions and recalculating the bias field, but also estimating the MRF parameters $\{\xi_{kk'}\}$ and $\{\nu_{kk'}\}$. As a result, the total iterative scheme now consists in four steps, shown in Fig. 1.11.

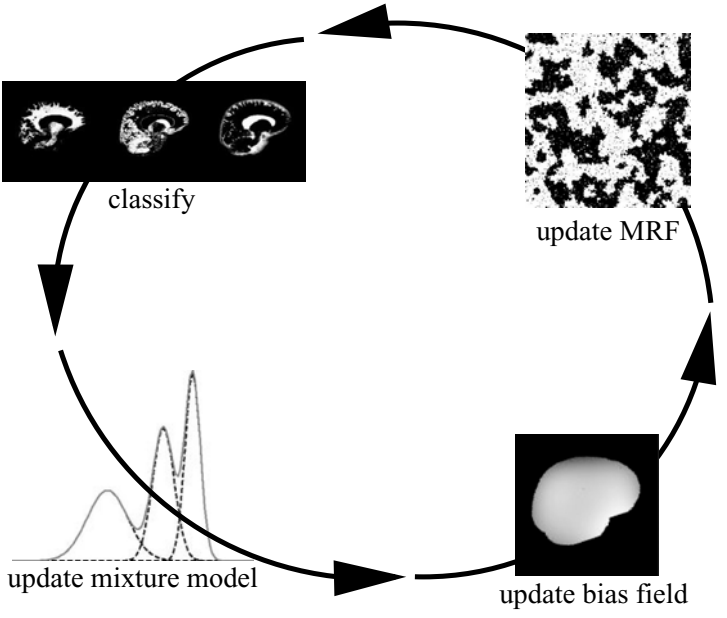
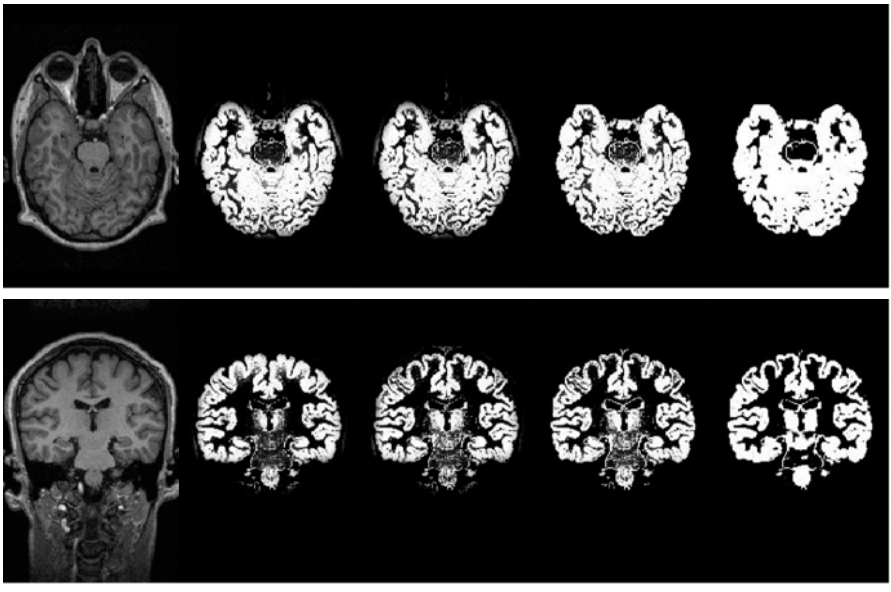


Figure 1.11: The extension of the model with a MRF prior results in a four-step algorithm that interleaves classification, estimation of the normal distributions, bias field correction, and estimation of the MRF parameters.

The calculation of the MRF parameters poses a difficult problem for which a heuristic, noniterative approach is used. For each neighborhood configuration $(\mathcal{N}^p, \mathcal{N}^o)$, the number of times that the central voxel belongs to class k in the current classification is compared to the number of times it belongs to class k', for every couple of classes (k, k'). This results in an overdetermined linear system of equations that is solved for the MRF parameters $(\xi_{kk'}, \nu_{kk'})$ using a least squares fit procedure [40].

1.3.2 Example

Figure 1.12 demonstrates the influence of each component of the algorithm on the resulting segmentations of a T1-weighted image. First, the method of section 1.2, where each voxel is classified independently, was used without the bias



(a) (b) (c) (d) (e)

Figure 1.12: Example of how the different components of the algorithm work. From left to right: (a) T1-weighted image, (b) gray matter segmentation without bias field correction and MRF, (c) gray matter segmentation with bias field correction but without MRF, (d) gray matter segmentation with bias field correction and MRF, and (e) gray matter segmentation with bias field correction and MRF without constraints. (Source: Ref. [39].)

correction step (Fig. 1.12(b)). It can be seen that white matter at the top of the brain is misclassified as gray matter. This was clearly improved when the bias field correction step was added (Fig. 1.12(c)). However, some tissues surrounding the brain have intensities that are similar to brain tissue and are wrongly classified as gray matter. With the MRF model described in section 1.3.1, a better distinction is obtained between brain tissues and tissues surrounding the brain (Fig. 1.12(d)). This is most beneficial in case of single-channel MR data, where it is often difficult to differentiate such tissues based only on their intensity. The MRF cleans up the segmentations of brain tissues, while preserving the detailed interface between gray and white matter, and between gray matter and CSF. Figure 1.13 depicts a 3-D volume rendering of the gray matter segmentation map when the MRF is used.

To demonstrate the effect of the constraints on the MRF parameters $\xi_{kk'}$ and $\nu_{kk'}$ described in section 1.3.1, the same image was processed without such constraints (Fig. 1.12(e)). The resulting segmentation shows nicely distinct regions, but small details, such as small ridges of white matter, are lost. The MRF prior has overregularized the segmentation and should therefore not be used in this form.

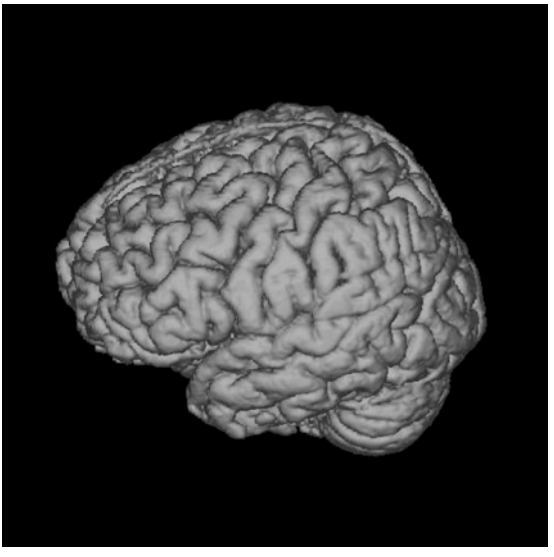


Figure 1.13: A 3-D volume rendering of the gray matter segmentation of the data of Fig. 1.12 with bias field correction and MRF. (Source: Ref. [39].)

1.3.3 Validation and Conclusions

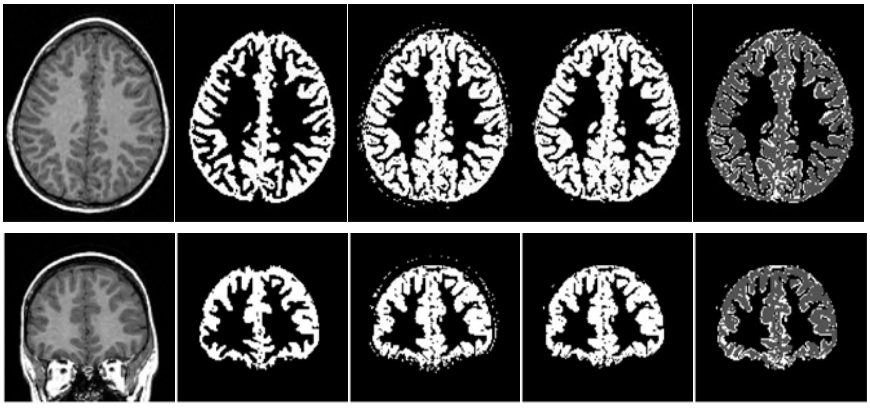
The method was validated on simulated MR images of the head that were generated by the BrainWeb MR simulator [45], for varying number of MR channels, noise, and severity of the bias fields. The automatic segmentations of each tissue k were compared with the known ground truth by calculating the similarity index

$$\frac{2V_{12}^k}{V_1^k + V_2^k} \tag{1.8}$$

where V_{12}^k denotes the volume of the voxels classified as tissue k by both raters, and V_1^k and V_2^k the volume of class k assessed by each of the raters separately. This metric, first described by Dice [46] and recently re-introduced by Zijdenbos *et al.* [47], attains the value of 1 if both segmentations are in full agreement, and 0 if there is no overlap at all. For all the simulated data, it was found that the total brain volume was accurately segmented, but the segmentation of gray matter and white matter individually did generally not attain the same accuracy. This was caused by misclassification of the white matter–gray matter interface, where PV voxels do not belong to either white matter or gray matter, but are really a mixture of both.

The automated method was also validated by comparing its segmentations of nine brain MR scans of children to the manual tracings by a human expert. The automated and manual segmentations showed an excellent similarity index of 95% on average for the total brain, but a more moderate similarity index of 83% for gray matter. Figure 1.14 depicts the location of misclassified gray matter voxels for a representative dataset. It can be seen that the automatic algorithm segments the gray matter–CSF interface in more detail than the manual tracer. Some tissue surrounding the brain is still misclassified as gray matter, although this error is already reduced compared to the situation where no MRF prior is used. However, by far most misclassifications are due to the classification of gray–white matter PV voxels to gray matter by the automated method. The human observer has segmented white matter consistently as a thicker structure than the automatic algorithm.

Finally, Park *et al.* [48] tested the method presented in this section on a database of T1-weighted MR scans of 20 normal subjects. These data along with expert manual segmentations are made publicly available on the WWW by the



(a) (b) (c) (d) (e)

Figure 1.14: Comparison between manual delineation and automated tissue classification on a representative dataset. From left to right: (a) axial and coronal slice, (b) corresponding manual segmentation of gray matter, (c) automatic segmentation of gray matter without MRF prior, (d) automatic segmentation of gray matter with MRF, and (e) difference between manual and automatic segmentation with MRF shown in white. (Source: Ref. [39].)

Center for Morphometric Analysis at the Massachusetts General Hospital.[4] The segmented image for each case was compared with the corresponding expert manual segmentation by means of an overlap metric. The overlap measure is simply the intersection of two coregistered images divided by their union. Thus, similar pairs of images approach an overlap value of 1, while dissimilar pairs approach 0. Overall, the method presented here outperformed the other listed automated methods (see [49] for details on these other methods) with an average overlap measure of 0.681 and 0.708 for GM and WM, respectively (compare to a more recent result: (0.662, 0.683) in [50]). It performed especially better on the difficult cases where there were significant bias fields associated with the image.

Partial voluming violates the model assumption that each voxel belongs to only one single class. In reality, PV voxels are a mixture of tissues and every segmentation method that tries to assign them exclusively to one class is doomed to fail. The problem is especially important in images of the brain since the interface between gray and white matter is highly complex, which results in a high

[4] http://www.cma.mgh.harvard.edu/ibsr.

volume of PV voxels compared to the volume of pure tissue voxels. Misclassification of this thin interface therefore gives immediate rise to considerable segmentation errors [51].

1.4 Model Outliers and Robust Parameter Estimation

So far, only the segmentation of MR images of normal brains has been addressed. In order to quantify MS lesions or other neuropathology related signal abnormalities in the images, the method needs to be further extended. Adding an explicit model for the pathological tissues is difficult because of the wide variety of their appearance in MR images, and because not every individual scan contains sufficient pathology for estimating the model parameters. These problems can be circumvented by detecting lesions as voxels that are not well explained by the statistical model for normal brain MR images [52]. The method presented here simultaneously detects the lesions as model outliers, and excludes these outliers from the model parameter estimation.

1.4.1 Background

Suppose that J samples $\boldsymbol{y}_j$, $j = 1, 2, \ldots, J$ are drawn independently from a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ that are grouped in $\Phi = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ for notational convenience. Given these samples, the ML parameters Φ can be assessed by maximizing

$$\sum_j \log f(\boldsymbol{y}_j \mid \Phi) \tag{1.9}$$

which yields

$$\begin{aligned} \boldsymbol{\mu} &= \frac{\sum_j \boldsymbol{y}_j}{J} \\ \boldsymbol{\Sigma} &= \frac{\sum_j (\boldsymbol{y}_j - \boldsymbol{\mu})(\boldsymbol{y}_j - \boldsymbol{\mu})^t}{J} \end{aligned} \tag{1.10}$$

In most practical applications, however, the assumed normal model is only an approximation to reality, and estimation of the model parameters Φ should not be severely affected by the presence of a limited amount of model outliers.

Considerable research efforts in the field of robust statistics [53] have resulted in a variety of methods for robust estimation of model parameters in the presence of outliers, from which the so-called M-estimators [53] present the most popular family.

Considering Eq. 1.9, it can be seen that the contribution to the loglikelihood of an observation that is atypical of the normal distribution is high, since $\lim_{f(\boldsymbol{y} \mid \Phi)\to 0} \log f(\boldsymbol{y} \mid \Phi) = -\infty$. The idea behind M-estimators is to alter Eq. (1.9) slightly in order to reduce the effect of outliers. A simple way to do this, which has recently become very popular in image processing [54] and medical image processing [6, 16, 17, 55], is to model a small fraction of the data as being drawn from a rejection class that is assumed to be uniformly distributed. It can be shown that assessing the ML parameters is now equivalent to maximizing

$$\sum_j \log(f(\boldsymbol{y}_j \mid \Phi) + \lambda), \quad \lambda \geq 0 \tag{1.11}$$

with respect to the parameters Φ, where λ is an *a priori* chosen threshold [54]. Since $\lim_{f(\boldsymbol{y} \mid \Phi)\to 0} \log(f(\boldsymbol{y} \mid \Phi) + \lambda) = \log(\lambda)$, the contribution of atypical observations on the log-likelihood is reduced compared to Eq. (1.9).

One possibility to numerically maximize Eq. (1.11) is to calculate iteratively the weights

$$t(\boldsymbol{y}_j \mid \Phi^{(m-1)}) = \frac{f(\boldsymbol{y}_j \mid \Phi^{(m-1)})}{f(\boldsymbol{y}_j \mid \Phi^{(m-1)}) + \lambda} \tag{1.12}$$

based on the parameter estimation $\Phi^{(m-1)}$ in iteration $(m-1)$, and subsequently update the parameters $\Phi^{(m)}$ accordingly:

$$\begin{aligned} \boldsymbol{\mu}^{(m)} &= \frac{\sum_j t(\boldsymbol{y}_j \mid \Phi^{(m-1)}) \cdot \boldsymbol{y}_j}{\sum_j t(\boldsymbol{y}_j \mid \Phi^{(m-1)})} \\ \boldsymbol{\Sigma}^{(m)} &= \frac{\sum_j t(\boldsymbol{y}_j \mid \Phi^{(m-1)}) \cdot (\boldsymbol{y}_j - \boldsymbol{\mu}^{(m)})(\boldsymbol{y}_j - \boldsymbol{\mu}^{(m)})^t}{\sum_j t(\boldsymbol{y}_j \mid \Phi^{(m-1)})} \end{aligned} \tag{1.13}$$

Solving an M-estimator by iteratively recalculating weights and updating the model parameters based on these weights is commonly referred to as the W-estimator [56]. The weight $t(\boldsymbol{y}_j \mid \Phi) \in [0, 1]$ reflects the typicality of sample i with respect to the normal distribution. For typical samples, $t(\boldsymbol{y}_j \mid \Phi) \simeq 1$, whereas $t(\boldsymbol{y}_j \mid \Phi) \simeq 0$ for samples that deviate far from the model. Comparing Eq. (1.13) with Eq. (1.10), it can therefore be seen that the M-estimator effectively

down-weights observations that are atypical for the normal distribution, making the parameter estimation more robust against such outliers.

1.4.2 From Typicality Weights to Outlier Belief Values

Since each voxel j has only a contribution of $t(\boldsymbol{y}_j \mid \Phi)$ to the parameter estimation, the remaining fraction

$$1 - t(\boldsymbol{y}_j \mid \Phi) \tag{1.14}$$

reflects the belief that it is a model outlier. The ultimate goal in our application is to identify these outliers as they are likely to indicate pathological tissues. However, the dependence of Eq. (1.14) through $t(\boldsymbol{y}_j \mid \Phi)$ on the determinant of the covariance matrix $\boldsymbol{\Sigma}$ prevents its direct interpretation as a true outlier belief value.

In statistics, an observation $\boldsymbol{y}$ is said to be abnormal with respect to a given normal distribution if its so-called Mahalanobis distance

$$d = \sqrt{(\boldsymbol{y} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\mu})}$$

exceeds a predefined threshold. Regarding Eq. (1.12), the Mahalanobis distance at which the belief that a voxel is an outlier exceeds the belief that it is a regular sample decreases with increasing $|\boldsymbol{\Sigma}|$. Therefore, the Mahalanobis distance threshold above which voxels are considered abnormal changes over the iterations as $\boldsymbol{\Sigma}$ is updated. Because of this problem, it is not clear how λ should be chosen.

Therefore, Eq. (1.12) is modified into

$$t(\boldsymbol{y}_j \mid \Phi^{(m-1)}) = \frac{f(\boldsymbol{y}_j \mid \Phi^{(m-1)})}{f(\boldsymbol{y}_j \mid \Phi^{(m-1)}) + \frac{1}{\sqrt{(2\pi)^C |\boldsymbol{\Sigma}^{(m-1)}|}} \exp(-\frac{1}{2}\kappa^2)}$$

where $|\boldsymbol{\Sigma}|$ is explicitly taken into account and where λ is reparameterized using the more easily interpretable κ. This $\kappa \geq 0$ is an explicit Mahalanobis distance threshold that specifies a statistical significance level, as illustrated in Fig. 1.15. The lower κ is chosen, the easier voxels are considered as outliers. On the other hand, choosing $\kappa = \infty$ results in $t(\boldsymbol{y}_j \mid \Phi^{(m-1)}) = 1, \forall \boldsymbol{j}$ which causes no outliers to be detected at all.

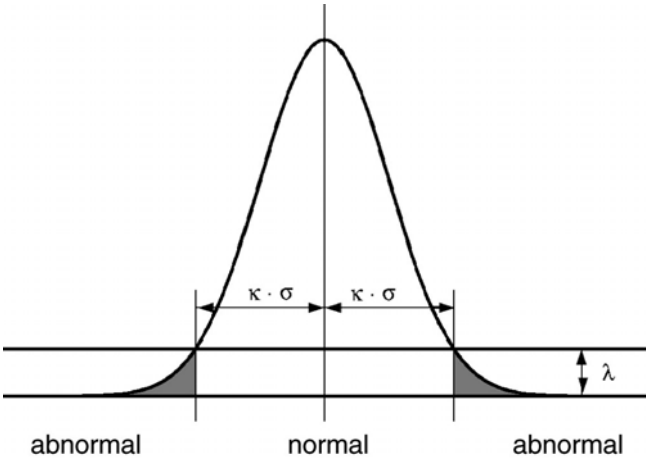


Figure 1.15: The threshold κ defines the Mahalanobis distance at which the belief that a voxel is a model outlier exceeds the belief that it is a regular sample (this figure depicts the unispectral case, where $\Sigma = \sigma^2$).

1.4.3 Robust Estimation of MR Model Parameters

Based on the same concepts, the EM framework used in the previous sections for estimating the parameters of models for normal brain MR images can be extended to detect model outliers such as MS lesions. In the original EM algorithm, a statistical classification $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)})$ is performed in the expectation step, and the subsequent maximization step involves updating the model parameters according to this classification. The weights $f(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)}), k = 1, 2, \ldots, K$ represent the degree to which voxel j belongs to each of the K tissues. However, since $\sum_k f(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)}) = 1$, an observation that is atypical for each of the normal distributions cannot have a small membership value for all tissue types simultaneously.

A similar approach as the one described above, where Eq. (1.9) was replaced with the more robust Eq. (1.11) and solved with a W-estimator, results in a maximization step in which model outliers are down-weighted. The resulting equations for updating the model parameters are identical to the original ones, provided that the weights $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)})$ are replaced everywhere with a combination of two weights $f(l_j \mid \boldsymbol{Y}, \Phi^{(m-1)}) \cdot t(\boldsymbol{y}_j \mid l_j, \Phi^{(m-1)})$, where

$$t(\boldsymbol{y}_j \mid l_j, \Phi^{(m-1)}) = \frac{f(\boldsymbol{y}_j \mid l_j, \Phi^{(m-1)})}{f(\boldsymbol{y}_j \mid l_j, \Phi^{(m-1)}) + \frac{1}{\sqrt{(2\pi)^c |\Sigma_\kappa^{(m-1)}|}} \exp(-\frac{1}{2}\kappa^2)} \tag{1.15}$$

reflects the degree of typicality of voxel j in tissue class l_j. Since $\sum_k f(l_j = k \mid \boldsymbol{Y}, \Phi^{(m-1)}) \cdot t(\boldsymbol{y}_j \mid l_j = k, \Phi^{(m-1)})$ is not constrained to be unity, model outliers can have a small degree of membership in all tissue classes simultaneously. Therefore, observations that are atypical for each of the K tissue types have a reduced weight on the parameter estimation, which robustizes the EM-procedure. Upon convergence of the algorithm, the belief that voxel j is a model outlier is given by

$$1 - \sum_k f(l_j = k \mid \boldsymbol{Y}, \Phi) \cdot t(\boldsymbol{y}_j \mid l_j = k, \Phi) \tag{1.16}$$

Section 1.5 discusses the use of this outlier detection scheme for fully automated segmentation of MS lesions from brain MR images.

1.5 Application to Multiple Sclerosis

In [52], the outlier detection scheme of section 1.4 was applied for fully automatic segmentation of MS lesions from brain MR scans that consist of T1-, T2-, and PD-weighted images. Unfortunately, outlier voxels also occur outside MS lesions. This is typically true for partial volume voxels that, in contravention to the assumptions made, do not belong to one single tissue type but are rather a mixture of more than one tissue. Since they are perfectly normal brain tissue, they are prevented from being detected as MS lesion by introducing constraints on intensity and context on the weights $t(\boldsymbol{y}_j \mid l_j, \Phi)$ calculated in Eq. (1.15).

1.5.1 Intensity and Contextual Constraints

- Since MS lesions appear hyperintense on both the PD- and the T2-weighted images, only voxels that are brighter than the mean intensity of gray matter in these channels are allowed to be outliers.

- Since around 90–95% of the MS lesions are white matter lesions, the contextual constraint is added that MS lesions should be located in the vicinity of white matter. In each iteraction, the normal white matter is fused with the lesions to form a mask of the total white matter. Using a MRF as in section 1.3, a voxel is discouraged from being classified as MS lesion in the absence of neighboring white matter. Since the MRF parameters are estimated from the data in each iteration as in section 1.3, these contextual constraints automatically adapt to the voxel size of the data.

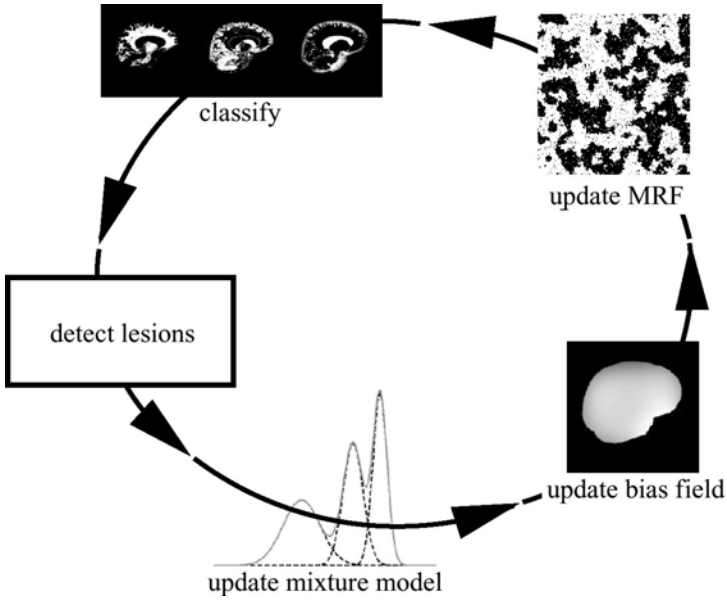


Figure 1.16: The complete method for MS lesion segmentation iteratively interleaves classification of the voxels into normal tissue types, MS lesion detection, estimation of the normal distributions, bias field correction, and MRF parameter estimation.

The complete method is summarized in Fig. 1.16. It iteratively interleaves statistical classification of the voxels into normal tissue types, assessment of the belief for each voxel that it is not part of an MS lesion based on its intensity and on the classification of its neighboring voxels, and, only based on what is considered as normal tissue, estimation of the MRF, intensity distributions, and bias field parameters. Upon convergence, the belief that voxel j is part of an MS lesion is obtained by Eq. (1.16). The method is fully automated, with only one single parameter that needs to be experimentally tuned: the Mahalanobis threshold κ in Eq. (1.15). A 3-D rendering of the segmentation maps including the segmentation fo MS lesions is shown in Fig. (1.17).

1.5.2 Validation

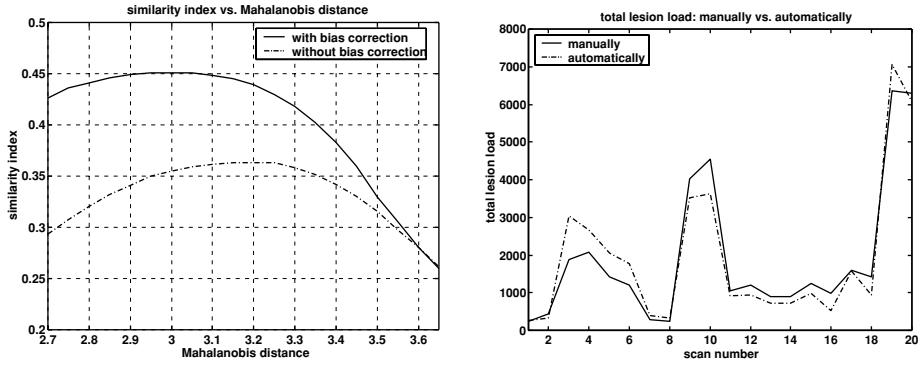
As part of the BIOMORPH project [57], we analyzed MR data acquired during a clinical trial in which 50 MS patients were repeatedly scanned with an interval of approximately 1 month over a period of about 1 year. The serial image data consisted at each time point of a PD/T2-weighted image pair and

(a) (b) (c)

Figure 1.17: A 3-D rendering of (a) gray matter, (b) white matter, and (c) MS lesion segmentation maps. (Color slide)

a T1-weighted image with 5 mm slice thickness. From 10 of the patients, two consecutive time points were manually analyzed by a human expert who traced MS lesions based only on the T2-weighted images. The automatic algorithm was repeatedly applied with values of the Mahalanobis distance κ varying from 2.7 (corresponding to a significance level of $p = 0.063$) to 3.65 (corresponding to $p = 0.004$), in steps of 1.05. The automatic delineations were compared with the expert segmentations by comparing the so-called total lesion load (TLL), measured as the number of voxels that were classified as MS lesion, on these 20 scans. The TLL value calculated by the automated method decreased when κ was increased, since the higher the κ, the less easily voxels are rejected from the model. Varying κ from 2.7 to 3.65 resulted in an automatic TLL of respectively 150% to only 25% of the expert TLL. However, despite the strong influence of κ on the absolute value of the TLL, the linear correlation between the automated TLLs of the 20 scans and the expert TLLs was remarkable insensitive to the choice of κ. Over this wide range, the correlation coefficient varied between 0.96 and 0.98.

Comparing the TLL of two raters does not take into account any spatial correspondence of the segmented lesions. We therefore calculated the similarity index defined in Eq. (1.8), which is simply the volume of intersection of the two segmentations divided by the mean of the two segmentation volumes. For the 20 scans, Fig. 1.18(a) depicts the value of this index for varying κ, both with and without the bias correction step included in the algorithm, clearly demonstrating the need for bias field correction. The best correspondence, with a similarity index of 0.45, was found for $\kappa \simeq 3$. For this value of κ, the automatic TLL was virtually equal to the expert TLL, and therefore, a similarity index of



(a) (b)

Figure 1.18: (a) Similarity index between the automatic and the expert lesion delineations on 20 images for varying κ, with and without the bias field correction component enabled in the automated method. (b) The 20 automatic total lesion load measurements for $\kappa = 3$ shown along with the expert measurements. (Source: Ref. [52].)

0.45 means that less than half of the voxels labeled as lesion by the expert were also identified by the automated method and vice versa.

For illustration purposes, the expert TLLs of the 20 scans are depicted along with the automatic ones for $\kappa = 3$ in Fig. 1.18(b). A paired t test did not reveal a significant difference between the manual and these automatic TLL ratings ($p = 0.94$). Scans 1 and 2 are two consecutive scans from one patient, 3 and 4 from the next and so on. Note that in 9 out of 10 cases, the two ratings agree over the direction of the change of the TLL over time. Figure 1.19 displays the MR data of what is called scan 19 in Fig. 1.18(b) and the automatically calculated classification along with the lesion delineations performed by the human expert.

1.5.3 Discussion

Most of the methods for MS lesion segmentation described in the literature are semiautomated rather than fully automated methods, designed to facilitate the tedious task of manually outlining lesions by human experts, and to reduce the inter- and intrarater variability associated with such expert segmentations. Typical examples of user interaction in these approaches include accepting or rejecting automatically computed lesions [58] or manually drawing regions of

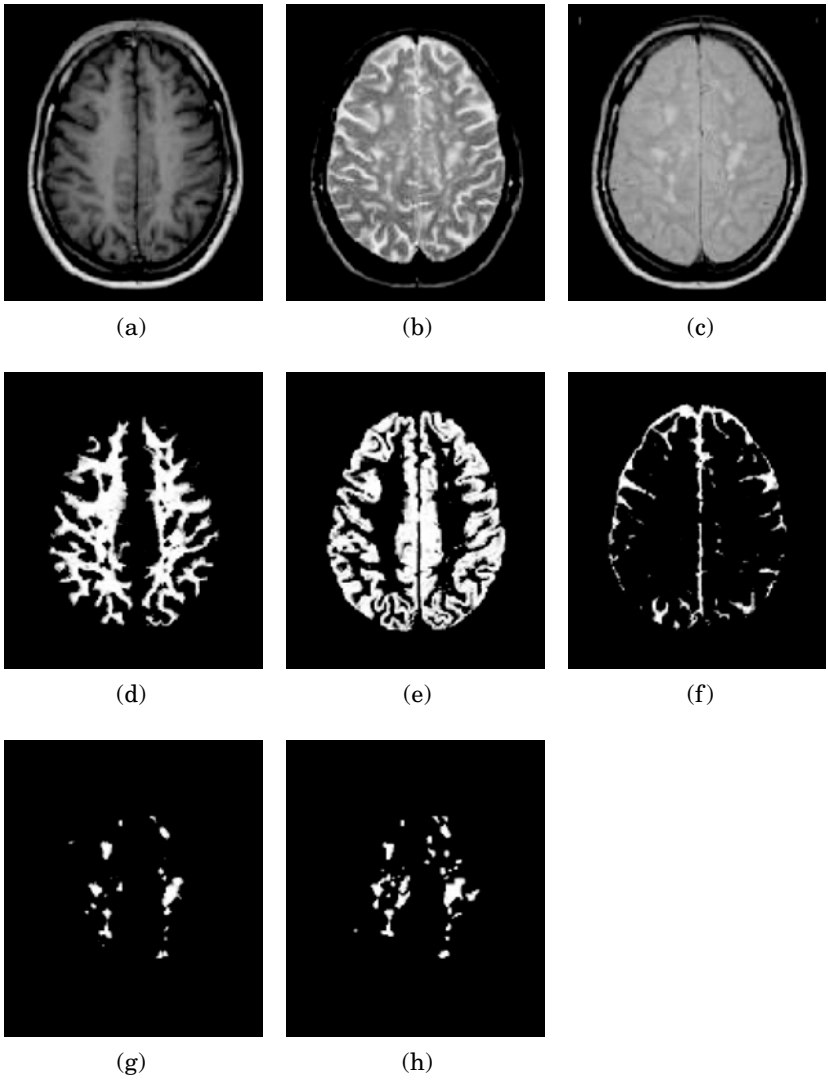


Figure 1.19: Automatic classification of one of the 20 serial datasets that were also analyzed by a human expert. (a) T1-weighted image; (b) T2-weighted image; (c) PD-weighted image; (d) white matter classification; (e) gray matter classification; (f) CSF classification; (g) MS lesion classification; (h) expert delineation of the MS lesions. (Source: Ref. [52].)

pure tissue types for training an automated classifier [58–61]. While these methods have proven to be useful, they remain impractical when hundreds of scans need to be analyzed as part of a clinical trial, and the variability of manual tracings is not totally removed. In contrast, the method presented here is fully automated as it uses a probabilistic brain atlas to train its classifier. Furthermore the atlas provides spatial information that avoids nonbrain voxels from being classified as MS lesion, making the method work without the often-used tracing of the intracranial cavity in a preprocessing step [58–63].

A unique feature of our algorithm is that it automatically adapts its intensity models and contextual constraints when analyzing images that were acquired with a different MR pulse sequence or voxel size. Zijdenbos *et al.* described [64] and validated [65] a fully automated pipeline for MS lesion segmentation based on an artificial neural network classifier. Similarly, Kikinis, Guttmann *et al.* [62, 66] have developed a method with minimal user intervention that is built on the EM classifier of Wells *et al.* [4] with dedicated pre- and postprocessing steps. Both methods use a fixed classifier that is trained only once and that is subsequently used to analyze hundreds of scans. In clinical trials, however, interscan variations in cluster shape and location in intensity space cannot be excluded, not only because of hardware fluctuations of MR scanners over a period of time, but also because different imagers may be used in a multicenter trial [66]. In contrast to the methods described above, our algorithm retrains its classifier on each individual scan, making it adaptive to such contrast variations.

Often, a postprocessing step is applied to automatically segmented MS lesions, in which false positives are removed based on a set of experimentally tuned morphologic operators, connectivity rules, size thresholds, etc. [59, 60, 62]. Since such rules largely depend on the voxel size, they may need to be retuned for images with a different voxel size. Alternatively, images can be resampled to a specific image grid before processing, but this introduces partial voluming that can reduce the detection of lesions considerably, especially for small lesion loads [66]. To avoid these problems, we have added explicit contextual constraints on the iterative MS lesions detection that automatically adapt to the voxel size. Similar to other methods [59, 61, 63, 64], we exploit the knowledge that the majority of MS lesions occurs inside white matter. Our method fuses the normal white matter with the lesions in each iteration, producing, in combination with MRF constraints, a prior probability mask for white matter that is automatically updated during the iterations. Since the MRF parameters are

reestimated for each individual scan, the contextual constraints automatically adapt to the voxel size of the images.

Although the algorithm we present is fully automatic, an appropriate Mahalanobis distance threshold κ has to be chosen in advance. When evaluating the role of κ, a distinction has to be made between the possible application areas of the method. In clinical trials, the main requirement for an automated method is that its measurements change in response to a treatment in a manner proportionate to manual measurements, rather than having an exact equivalence in the measurements [67, 68]. In section 1.5.2 it was shown that the automatic measurements always kept changing proportionately to the manual measurements for a wide range of κ, with high correlation coefficients between 0.96 and 0.98. Therefore, the actual choice of κ is fairly unimportant for this type of application. However, the role of κ is much more critical when the goal is to investigate the basic MS mechanisms or time correlations of lesion groups in MS time series, as these applications require that the lesions are also spatially correctly detected. In general, the higher the resolution and the better the contrast between lesions and unaffected tissue in the images, the easier MS lesions are detected by the automatic algorithm and the higher κ should be chosen. Therefore, the algorithm presumably needs to be tuned for different studies, despite the automatic adaptation of the tissue models and the MRF parameters to the data.

1.6 Application to Epilepsy

Epilepsy is the most frequent serious primary neurological illness. Around 30% of the epilepsy patients have epileptic seizures that are not controlled with medication. Epilepsy surgery is by far the most effective treatment for these patients. The aim of any presurgical evaluation in epilepsy surgery is to delineate the epileptogenic zone as accurate as possible. The epileptogenic zone is that part of the brain that has to be removed surgically in order to render the patient seizure-free.

We applied the framework presented in this chapter to detect structural abnormalities related to focal cortical dysplasia (FCD) epileptic lesions in the cortical and subcortical grey matter in high-resolution MR images of the human brain. FCD is characterized by a focal thickening of the cerebral cortex, loss of definition between the gray and the white matter at the site of the lesion,

and a hypointense T1-weighted MR signal in the gray matter. The approach is volumetric: feature images isomorphic to the original MR image are generated, representing the spatial distribution of grey matter density or, following Antel *et al.* [69, 70], related features such as the ratio of cortical thickness over local intensity gradient. Since these feature images show consistently thick regions in certain parts of the normal anatomy (e.g. cerebellum), the specificity (reduction of the number of false responses) of intrasubject detection of epileptogenic lesions can be increased by comparing the feature response images of patients with that of a group of normal controls. We used the machinery of statistical parametric mapping (SPM) [71], as standard in functional imaging studies, to make voxel-specific inferences.

First, each 3-D MR image is automatically segmented (using the method presented in this chapter) into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF), resulting in an image representing the individual spatial distribution of GM. The statistical priors (Fig. 1.3) for each tissue class are warped to each subject using a nonrigid multimodal free-form (involving many degrees of freedom) registration algorithm [24]. Segmentation using a combination of intensity-based tissue classification and geometry-based atlas registration helps in reducing the misclassification of extra-cerebral tissues as gray matter and aids in the reduction of false positive findings during the statistical analysis. The gray and white matter continuous classification are binarized by deterministically assigning each voxel to the class for which it has a maximum probability of occupancy among all classes considered. Next, we estimate the cortical thickness by the method described in [72]. The method essentially solves an Eikonal PDE in the domain of the gray matter and gives the cortical thickness at each GM voxel as the sum of the minimum euclidean distance of the voxel to the GM–WM interface and the GM–CSF interface. Following the method of [69, 70], we generate the feature maps, for each subject (normals and patients), by dividing the cortical thickness by the signal gradient in the gray matter region. Next, each feature image is geometrically normalized to a standard space using a 12 parameter affine MMI registration [21]. Subsequently, individual feature maps of patients are compared to those of 64 control subjects in order to detect statistically significant clusters of abnormal feature map values.

Figure 1.20 shows three orthogonal views of overlays of clusters of significantly abnormal grey matter voxels in a 3-D MR image of a single epileptic patient.

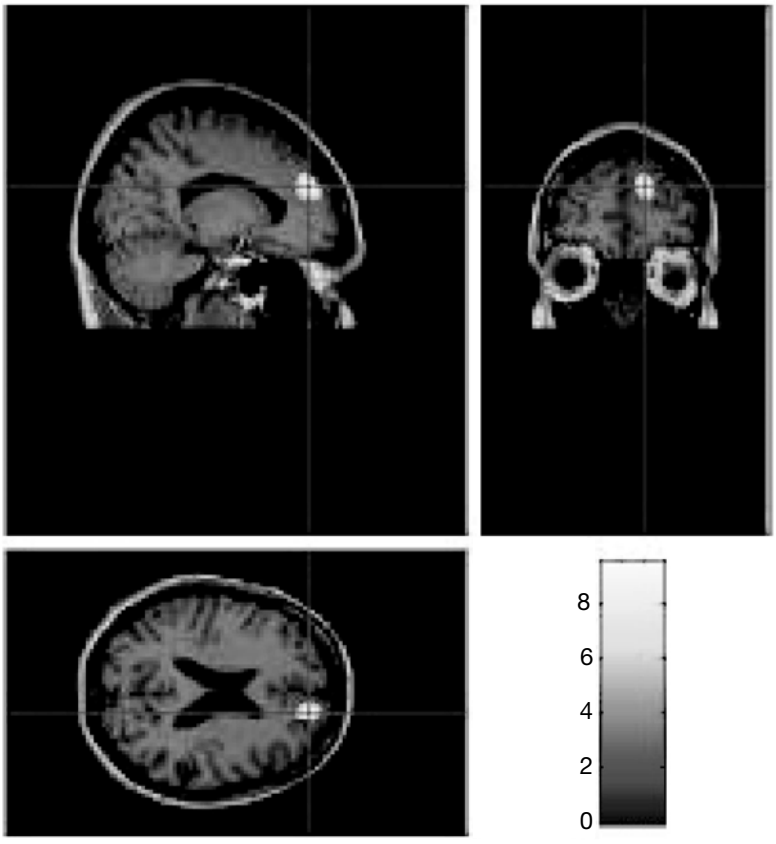


Figure 1.20: Cross-sectional overlay of the detected focal cortical thickening locus. The colour scale increases with statistical significance. Significance is measures by comparison of feature measurements (cortical thickness over intensity gradient) to a control subject database (Color Slide).

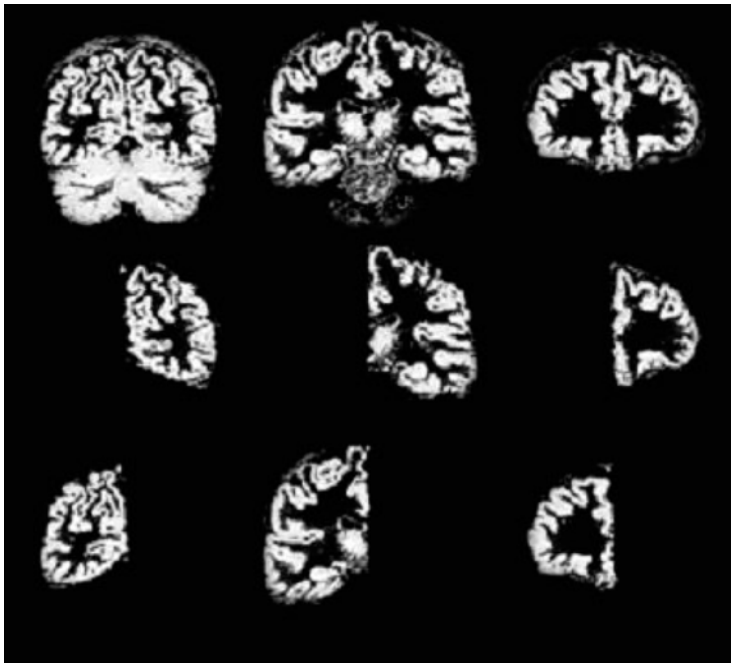
1.7 Application to Schizophrenia

Several studies have reported morphological differences in the brains of schizophrenic patients when compared to normal controls [73], such as enlargement of the cerebral ventricles and decreased or reversed cerebral asymmetries. These findings suggest the presence of structural brain pathology in schizophrenia. Some hypotheses have been proposed about schizophrenia as a syndrome of anomalies in normal brain development [74] whose origin may be genetically determined. Characterization of the morphological processes in schizophrenia

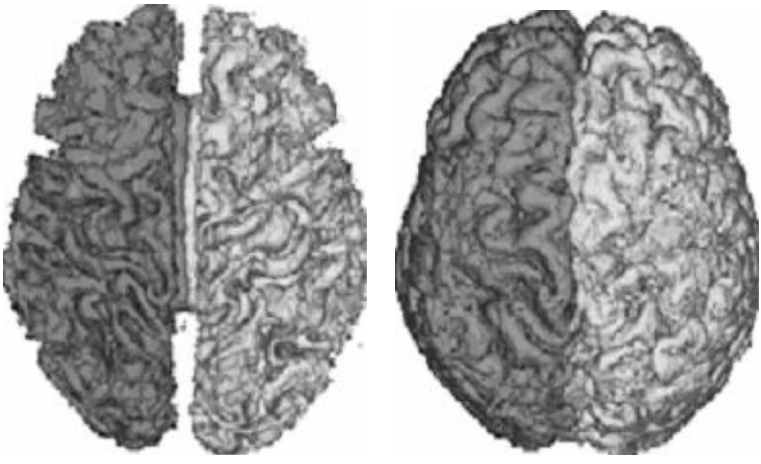
may lead to new pharmacological treatments aimed at prevention and cure rather than suppression of symptoms. There is in addition a particular focus on asymmetry as the defining characteristic of the human brain and the correlate of language. Techniques are required for describing and quantifying cerebral asymmetries to determine where these are located, how variable they are between individuals, and how the distribution in individuals with schizophrenia differs from that in the general population.

As an application of the framework proposed in this chapter, we present here a method for fully automatic quantification of cerebral grey and white matter asymmetry from MR images. White and grey matter are segmented after bias correction by the intensity-based tissue classification algorithm presented in section 1.2. Separation of the computed white and grey matter probability maps into left and right halves is achieved by nonrigid registration of the study image to a template MR image in which left and right hemispheres have been carefully segmented. The delineations of left and right hemispheres in the template image were transformed into binary label images, which were subsequently edited by morphological operations to match the brain envelope rather than the individual gyri and sulci to be more robust against differences in local cortical topology. The template image is matched to the study image by a combination of affine [21] and locally nonrigid transformations [24]. The resulting deformation field is then applied to the label images, such that matched outlines of left and right hemispheres are obtained. These are used to separate left and right halves in the original grey and white matter segmentations and, at the same time, to remove nonrelevant structures such as the cerebellum and brain stem. Finally, volumes for grey and white matter for each half of the brain separately are computed by integrating the corresponding probability maps within the brain regions of interest defined by the matched template image. Figure 1.21 illustrates that the grey and white matter segmentation maps obtained from the original images are correctly split in separate maps for left and right hemispheres by nonrigid registration with the labelled template image.

Various authors have presented alternative techniques to separate the brain hemispheres by the so-called midsagittal plane, defined as the plane that best fits the interhemispheric fissure of the brain [76] or as the plane that maximizes similarity between the image and its reflection relative to this plane [77, 78]. The advantage of the approach of intensity-driven non-rigid registration to a labelled template image, as presented here, is that it does not assume the boundary



(a)



(b) (c)

Figure 1.21: Left–right hemisphere grey matter and white matter separation on high quality high resolution data, consisting of a single sagittal T1-weighted image (Siemens Vision 1.5 T scanner, 3D MPRAGE, 256*256 matrix, 1.25 mm slice thickness, 128 slices, FOV = 256 mm, TR = 11.4 ms, TE = 4.4 ms) with good contrast between grey matter, white matter, and the tissues surrounding the brain. (a): Coronal sections through the grey matter segmentation map before (top row) and after left–right and cerebrum–cerebellum separation. (b) and (c): 3D Rendering of the segmented white and grey matter respectively. (Source: Ref. [75].)

between both hemispheres to be planar. We refer the reader to [75] for further details and discussion of the results obtained.

1.8 Conclusion

The model-based brain tissue classification framework presented here was setup to analyze MR signal abnormalities in neuropathological disorders in large sets of multispectral MR data in a reproducible and fully automatic way. The overall strategy adopted was to build statistical models for normal brain MR images, with emphasis on accurate intensity models. Signal abnormalities are detected as model outliers, i.e., voxels that cannot be well explained by the model. Special attention has been paid to automatically estimate all model parameters from the data itself to eliminate subjective manual tuning and training.

As discussed in section 1.1.1, geometry-driven and intensity-driven methods are the two main paradigms for model-based segmentation. In this chapter, both paradigms are used in a sequential order. Complex intensity models are developed that automatically fit to the data. As a result, multispectral MR data are segmented fully automatically without prior knowledge about the appearance of the different tissue types in the images. Bias fields are automatically corrected for, and the partial volume effect is explicitly taken into account. The intensity models are complemented with image-based brain atlas models of prior tissue probabilities. These atlases are first iconically matched to the images and subsequently used as constraints during the tissue classifications. The matching was originally limited to affine transformations between atlas and study image (see also [79], but was later extended to nonrigid transformations as well (see also [14, 50, 75, 80]).

These attempts to combine the ability of intensity-driven methods to capture local shape variations with the general description of global anatomy provided by geometry-driven methods have been limited to a sequential use of both methods in separate processing steps. Atlas maps of prior distributions of the different tissue classes are first geometrically aligned to the images to be segmented. These transformed maps provide an initial approximate segmentation to initialize the classification algorithm but also provide an estimate of the prior class probabilities for each voxel during further iterations.

In [13] an attempt was made to intertwine statistical intensity-based tissue classification and nonlinear registration of a digital anatomical template to segment both normal and abnormal anatomy. The algorithm iterates between a classification step to identify tissues and an elastic matching step to align a template of normal anatomy with the classified tissues. The alignment of the anatomical template is used to modify the classification to produce a spatially varying, rather than a global classification. The steps are iterated until the matched anatomy and the classification agree. However, this method currently needs manual supervision, and it still needs to be investigated how reliably this method can be automated. Moreover, the iterative procedure is not derived as the solution of an optimization problem. As a result, there is no guarantee that convergence, if at all, to a plausible solution can be obtained.

Wyatt and Noble [81], on the other hand, suggested a joint solution to the linked processes of segmentation and registration. They cast this as a maximum a posteriori (MAP) estimation of the segmentation labels and the geometric transformation parameters and pose the solution using MRF. Their results indicate that the addition of spatial priors (in the form of intermediate segmentation maps) leads to substantially greater robustness in rigid registration and the combination of data via registration improves the segmentation accuracy. However, their formulation is poorly suited for generalization to nonrigid registration.

D'Agostino *et al.* [82] explored the possibility of nonrigid image registration by maximizing an information theoretic measure of the similarity of voxel object labels directly, rather than of voxel intensities. Applied to intersubject MR brain image matching, such labels are obtained by the intensity-based tissue segmentation presented in this chapter, assigning each voxel a probability to belong to a particular tissue class. Using class labels as features for nonrigid image registration opens perspectives for integrating registration and segmentation as two cooperative processes in a single framework, by considering one of the images as an atlas that is nonrigidly warped onto the other and that provides *a priori* tissue distribution maps to guide the segmentation of the other image. The possibilities of such a method are enormous, since it would allow fully automated partial volume segmentation and bias correction of multispectral MR images with unknown tissue contrast, while deforming a label atlas at the same time. The quantification of intensity abnormalities could be confined to anatomical regions of interest. The brain could be automatically segmented into relevant substructures, allowing the quantification of changes in shape and volume, over

time in one individual patient, or between populations. Knowledge of the deformation of the label atlas would allow nonrigid multimodal registration of images of different patients and provide a common reference frame for population studies. Deriving realistic statistical models for the shape of the human brain is, therefore, a major challenge for further research.

Questions

1. *What is the parametric model of the MR bias field proposed in this chapter?*
2. *How are the parameters of the MR bias field estimated?*
3. *What are alternative models for the MR bias field?*
4. *How is the atlas-based geometric prior constructed?*
5. *How is the atlas of priors integrated into the classification?*
6. *What are the drawbacks of using such an atlas of priors and how to deal with it?*
7. *What is the Markov random field model used in brain tissue classification?*
8. *How do we limit the overregularization of Markov random fields?*

Bibliography

[1] Santago, P. and Gage, H., Quantification of MR brain images by mixture density and partial volume modeling, IEEE Trans. Med. Imaging, Vol. 12, No. 3, pp. 566–574, 1993.

[2] Laidlaw, D. H., Fleischer, K. W., and Barr, A. H., Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms, IEEE Trans. Med. Imaging, Vol. 17, No. 1, pp. 74–86, 1998.

[3] Choi, H. S., Haynor, D. R., and Kim, Y., Partial volume tissue classification of multichannel magnetic resonance images—A mixel model, IEEE Trans. Med. Imaging, Vol. 10, No. 3, pp. 395–407, 1991.

[4] Wells, W., III, Grimson, W., Kikinis, R., and Jolesz, F., Adaptive segmentation of MRI data, IEEE Trans. Med. Imaging, Vol. 15, No. 4, pp. 429–442, 1996.

[5] Held, K., Kops, E. R., Krause, B. J., Wells, W. M., III, Kikinis, R., and Müller-Gärtner, H. W., Markov random field segmentation of brain MR images, IEEE Trans. Med. Imaging, Vol. 16, No. 6, pp. 878–886, 1997.

[6] Guillemaud, R. and Brady, M., Estimating the bias field of MR Images, IEEE Trans. Med. Imaging, Vol. 16, No. 3, pp. 238–251, 1997.

[7] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput Vision, Vol. 1, No. 4, pp. 321–331, 1988.

[8] McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A survey, Med. Image Anal., Vol. 2, No. 1, pp. 1–36, 1996.

[9] Lötjönen, J., Reissman, P.-J., Mangin, I., and Katila, T., Model extraction from magnetic resonance volume data using the deformable pyramid, Med. Image Anal., Vol. 3, No. 4, pp. 387–406, 1999.

[10] Zeng, X., Staib, L., Schultz, R., and Duncan, J., Segmentation and measurement of the cortex from 3D MR images using coupled surfaces propagation, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 927–937, 1999.

[11] González Ballester, M., Zisserman, A., and Brady, M., Segmentation and measurement of brain structures in MRI including confidence bounds, Med. Image Anal., Vol. 4, pp. 189–200, 2000.

[12] Xu, C., Pham, D., Rettmann, M., Yu, D., and Prince, J., Reconstruction of the human cerebral cortex from magnetic resonance images, IEEE Trans. Med. Imaging, Vol. 18, No. 6, pp. 467–480, 1999.

[13] Warfield, S., Kaus, M., Jolesz, F., and Kikinis, R., Adaptive, template moderated, spatially varying statistical classification, Med. Image Anal., Vol. 4, No. 1, pp. 43–55, 2000.

[14] Collins, D. L., Zijdenbos, A. P., Barr, W. F. C., and Evans, A. C., ANIMAL+INSECT: Improved cortical structure segmentation, In: Proceedings of the Annual Symposium on Information Processing in Medical Imaging, Kuba, A., Samal, M., and Todd-Pokropek, A., eds., Lecture Notes in Computer Science, Vol. 1613, Springer, Berlin, pp. 210–223, 1999.

[15] Liang, Z., MacFall, J. R., and Harrington, D. P., Parameter estimation and tissue segmentation from multispectral MR images, IEEE Trans. Med. Imaging, Vol. 13, No. 3, pp. 441–449, 1994.

[16] Schroeter, P., Vesin, J.-M., Langenberger, T., and Meuli, R., Robust parameter estimation of intensity distributions for brain magnetic resonance images, IEEE Trans. Med. Imaging, Vol. 17, No. 2, pp. 172–186, 1998.

[17] Wilson, D. and Noble, J., An adaptive segmentation algorithm for time-of-flight MRA data, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 938–945, 1999.

[18] Dempster, A. P., Laird, N. M., and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc., Vol. 39, pp. 1–38, 1977.

[19] Wu, Z., Chung, H.-W., and Wehrli, F., A Bayesian approach to subvoxel tissue classification in NMR microscopic images of trabecular bone, MRM, Vol. 31, pp. 302–308, 1994.

[20] Evans, A., Collins, D., Mills, S., Brown, E., Kelly, R., and Peters, T., 3D statistical neuroanatomical models from 305 MRI volumes, In: Proceeding of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 1813–1817, 1993.

[21] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multi-modality image registration by maximization of mutual information, IEEE Trans. Med. Imaging, Vol. 16, No. 2, pp. 187–198, 1997.

[22] Maes, F., Vandermeulen, D., and Suetens, P., Medical image registration using mutual information, Proc. IEEE, Vol. 91, No. 10, pp. 1699–1722, 2003.

[23] Van Leemput, K., Maes, F., Vandermeulen, D., and Suetens, P., Automated model-based bias field correction of MR images of the brain, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 885–896, 1999.

[24] D'Agostino, E., Maes, F., Vandermeulen, D., and Suetens, P., A viscous fluid model for multimodal non-rigid image registration using mutual information, Med. Image Anal., Vol. 7, No. 4, pp. 565–575, 2003.

[25] Simmons, A., Tofts, P., Barker, G., and Arridge, S., Sources of intensity nonuniformity in spin echo images at 1.5 T, Magn. Reson. Med., Vol. 32, pp. 121–128, 1994.

[26] Sled, J. G. and Pike, G. B., Understanding Intensity Non-Uniformity in MRI, In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI'98, Lecture Notes in Computer Science, Vol. 1496, Springer, Berlin, pp. 614–622, 1998.

[27] Van Leemput, K., Maes, F., Vandermeulen, D., and Suetens, P., A unifying framework for partial volume segmentation of brain MR images, IEEE Trans. Med. Imaging, Vol. 22, No. 1, pp. 105–119, 2003.

[28] Tincher, M., Meyer, C., Gupta, R., and Williams, D., Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI, IEEE Trans. Med. Imaging, Vol. 12, No. 2, pp. 361–365, 1993.

[29] Moyher, S. E., Vigneron, D. B., and Nelson, S. J., Surface coil MR imaging of the human brain with an analytic reception profile correction, J. Magn. Reson. Imaging, Vol. 5, No. 2, pp. 139–144, 1995.

[30] González Ballester, M. A., Morphometric Analysis of Brain Structures in MRI, Ph.D. Thesis, Department of Engineering Science, University of Oxford, 1999.

[31] Dawant, B. M., Zijdenbos, A. P., and Margolin, R., Correction of intensity variations in MR images for computer-aided tissue classification, IEEE Trans. Med. Imaging, Vol. 12, No. 4, pp. 770–781, 1993.

[32] Meyer, C., Bland, P., and Pipe, J., Retrospective correction of MRI amplitude inhomogeneities, In: Proceedings of the First International Conference on Computer Vision, Virtual Reality, and Robotics in Medicine, CVRMED'95, Ayache, N., ed., Lecture Notes in Computer Science, Vol. 905, Springer, Nice, France, pp. 513–522, 1995.

[33] Brechbühler, C., Gerig, G., and Székely, G., Compensation of spatial inhomogeneity in MRI based on a parametric bias estimate, In: Proceedings of Visualization in Biomedical Computing, VBC'96, Lecture Notes in Computer Science, Vol. 1131, Springer, Berlin, pp. 141–146, 1996.

[34] Sled, J. G., Zijdenbos, A. P., and Evans, A. C., A comparison of retrospective intensity non-uniformity correction methods for MRI, In: Proceedings of XVth International Conference on Information Processing in Medical Imaging, IPMI'97, Lecture Notes in Computer Science, Vol. 1230, Springer, Berlin, pp. 459–464, 1997.

[35] Styner, M., Brechbühler, C., Székely, G., and Gerig, G., Parametric estimate of intensity inhomogeneities applied to MRI, IEEE Trans. Med. Imaging, Vol. 19, No. 3, pp. 153–165, 2000.

[36] Sled, J. G., Zijdenbos, A. P., and Evans, A. C., A Nonparametric method for automatic correction of intensity nonuniformity in MRI Data, IEEE Trans. Med. Imaging, Vol. 17, No. 1, pp. 87–97, 1998.

[37] Mangin, J.-F., Entropy minimization for automatic correction of intensity nonuniformity, In: Proceedings of IEEE Workshop on Mathematical

Methods in Biomedical Image Analysis, MMBIA'00, pp. 162–169, 2000.

[38] Likar, B., Viergever, M., and Pernus, F., Retrospective correction of MR intensity inhomogeneity by information minimization, In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI 2000, Lecture Notes in Computer Science, Vol. 1935, Springer, Berlin, pp. 375–384, 2000.

[39] Van Leemput, K., Maes, F., Vandermeulen, D., and Suetens, P., Automated model-based tissue classification of MR images of the brain, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 897–908, 1999.

[40] Li, S., Markov Random Field Modeling in Computer Vision, Computer Science Workbench Series, Springer, Berlin, 1995.

[41] Ising, E., Beitrag zur theorie des ferromagnetismus, Zeitschrift für Physik, Vol. 31, pp. 253–258, 1925.

[42] Descombes, X., Mangin, J.-F., Pechersky, E., and Sigelle, M., Fine structure preserving markov model for image processing, In: Proceedings of the 9th Scandinavian Conference on Image Analysis, SCIA'95, pp. 349–356, 1995.

[43] Zhang, J., The mean-field theory in EM procedures for Markov random fields, IEEE Trans. Signal Process., Vol. 40, No. 10, pp. 2570–2583, 1992.

[44] Langan, D. A., Molnar, K. J., Modestino, J. W., and Zhang, J., Use of the mean-field approximation in an EM-based approach to unsupervised stochastic model-based image segmentation, In: Proceedings of ICASSP'92, San Fransisco, CA, March 1992, Vol. 3, pp. 57–60.

[45] Kwan, R. K.-S., Evans, A. C., and Pike, G. B., MRI simulation-based evaluation of image-processing and classification methods, IEEE Trans. Med. Imaging, Vol. 18, No. 11, pp. 1085–1097, 1999. Available at http://www.bic.mni.mcgill.ca/brainweb/.

[46] Dice, L. R., Measures of the amount of ecologic association between species, Ecology, Vol. 26, No. 3, pp. 297–302, 1945.

[47] Zijdenbos, A. P., Dawant, B. M., and Margolin, R. A., Intensity correction and its effect on measurement variability in the computer-aided analysis of MRI, In: Proceedings of 9th International Symposium and Exhibition on Computer Assisted Radiology, CAR'95, Springer, Berlin, pp. 216–221, June 1995.

[48] Park, J., Gerig, G., Chakos, M., Vandermeulen, D., and Lieberman, J., Neuroimaging of psychiatry disease: Reliable and efficient automatic brain tissue segmentation for increased sensitivity, Schizophrenia Res., Vol. 49, p. 163, 1994.

[49] Rajapakse, J. and Krugge, F., Segmentation of MR images with intensity inhomogeneities, Image Vision Comput., Vol. 16, pp. 165–180, 1998.

[50] Marroquin, J. L., Vemuri, B. C., Botello, S., Calderon, F., and Fernandez-Bouzas, A., An accurate and efficient Bayesian method for automatic segmentation of brain MRI, IEEE Trans. Med. Imaging, Vol. 21, No. 8, pp. 934–945, 2002.

[51] Niessen, W., Vincken, K., Weickert, J., ter Haar Romeny, B., and Viergever, M., Multiscale segmentation of three-dimensional MR brain images, Int. J. Comput. Vision, Vol. 31, No. 2/3, pp. 185–202, 1999.

[52] Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., and Suetens, P., Automated segmentation of multiple sclerosis lesions by model outlier detection, IEEE Trans. Med. Imaging, Vol. 20, No. 8, pp. 677–688, 2001.

[53] Huber, P., Robust Statistics, Wiley series in Probability and Mathematical Statistics, Wiley, New York, 1981.

[54] Zhuang, X., Huang, Y., Palaniappan, K., and Zhao, Y., Gaussian mixture density modeling, decomposition, and applications, IEEE Trans. Image Process., Vol. 5, No. 9, pp. 1293–1302, 1996.

[55] Chung, A. and Noble, J., Statistical 3D vessel segmentation using a rician distribution, In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI'99, Lecture Notes in Computer Science, Vol. 1679, Springer, Berlin, pp. 82–89, 1999.

[56] Hoaglin, D., Mosteller, F., and Tukey, J., eds., Understanding Robust and Explanatory Data Analysis, Wiley series in Probability and Mathematical Statistics, Wiley, New York, 1983.

[57] European project on Brain Morphometry, BIOMORPH, EU-BIOMED2 Project No. BMH4-CT96-0845, 1996–1998.

[58] Udupa, J., Wei, L., Samarasekera, S., Miki, Y., van Buchem, M., and Grossman, R., Multiple sclerosis lesion quantification using fuzzy-connectedness principles, IEEE Trans. Med. Imaging, Vol. 16, No. 5, pp. 598–609, 1997.

[59] Johnston, B., Atkins, M., Mackiewich, B., and Anderson, M., Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI, IEEE Trans. Med. Imaging, Vol. 15, No. 2, pp. 154–169, 1996.

[60] Zijdenbos, A., Dawant, B. M., Margolin, R. A., and Palmer, A. C., Morphometric analysis of white matter lesions in MR images: Method and validation, IEEE Trans. Med. Imaging, Vol. 13, No. 4, pp. 716–724, 1994.

[61] Kamber, M., Shinghal, R., Collins, D., Francis, G., and Evans, A., Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images, IEEE Trans. Med. Imaging, Vol. 14, No. 3, pp. 442–453, 1995.

[62] Kikinis, R., Guttmann, C., Metcalf, D., Wells, W., III, Ettinger, G., Weiner, H., and Jolesz, F., Quantitative follow-up of patients with multiple sclerosis using MRI: Technical aspects, J. Magn. Reson. Imaging, Vol. 9, No. 4, pp. 519–530, 1999.

[63] Warfield, S., Dengler, J., Zaers, J., Guttmann, C., Wells, W., III, Ettinger, G., Hiller, J., and Kikinis, R., Automatic identification of grey matter structures from MRI to improve the segmentation of white matter lesions, J. Image Guided Surg., Vol. 1, No. 6, pp. 326–338, 1995.

[64] Zijdenbos, A., Evans, A., Riahi, F., Sled, J., Chui, J., and Kollokian, V., Automatic quantification of multiple sclerosis lesion volume using stereotaxic space, In: Proceedings of Visualization in Biomedical Computing, VBC'96, Lecture Notes in Computer Science, Springer, Berlin, Vol. 1131, pp. 439–448, 1996.

[65] Zijdenbos, A., Forghani, R., and Evans, A., Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT, In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI'98, Lecture Notes in Computer Science, Vol. 1496, Springer, Berlin, pp. 439–448, 1998.

[66] Guttmann, C., Kikinis, R., Anderson, M., Jakab, M., Warfield, S., Killiany, R., Weiner, H., and Jolesz, F., Quantitative follow-up of patients with multiple sclerosis using MRI: Reproducibility, J. Magn. Reson. Imaging, Vol. 9, No. 4, pp. 509–518, 1999.

[67] Evans, A., Frank, J., Antel, J., and Miller, D., The role of MRI in clinical trials of multiple sclerosis: Comparison of image processing techniques, Ann. Neurol., Vol. 41, No. 1, pp. 125–132, 1997.

[68] Filippi, M., Horsfield, M., Tofts, P., Barkhof, F., Thompson, A., and Miller, D., Quantitative assessment of MRI lesion load in monitoring the evolution of multiple sclerosis, Brain, Vol. 118, pp. 1601–1612, 1995.

[69] Antel, S. B., Bernasconi, A., Bernasconi, N., Collins, D. L., Kearney, R. E., Shinghal, R., and Arnold, D. L., Computational models of MRI characteristics of focal cortical dyaplasia improve lesion detection, NeuroImage, Vol. 17, No. 4, pp. 1755–1760, 2002.

[70] Antel, S. B., Collins, D. L., Bernasconi, N., Andermann, F., Shinghal, R., Kearney, R. E., Arnold, D. L., and Bernasconi, A., Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis, NeuroImage, Vol. 19, No. 4, pp. 1784–1759, 2003.

[71] Ashburner, J., Friston, K., Holmes, A., and Poline, J.-B., Statistical Parametric Mapping, The Wellcome Department of Cognitive Neurology, University College London, London. Available at http://www.fil.ion.ucl.ac.uk/spm/.

[72] Srivastava, S., Vandermeulen, D., Maes, F., Dupont, P., van Paesschen, W., and Suetens, P., An automated 3D algorithm for neo-cortical thickness measurement, In: Proceedings of Medical Image Computing and

Computer-Assisted Intervention, MICCAI'03, Lecture Notes in Computer Science, Springer, Berlin, Vol. 2879, pp. 488–495.

[73] DeLisi, L., Tew, W., Shu-Hong, X., Hoff, A. L., Sakuma, M., Kushner, M., Lee, G., Shedlack, K., Smith, A. M., and Grimson, R., A prospective follow-up study of brain morphology and cognition in first-epsiode schizophrenic patients: Preliminary findings, Biol. Psychiatry, Vol. 38, No. 2, pp. 349–360, 1995.

[74] Crow, T., Ball, J., Bloom, S., Brown, R., Bruton, C. J., Colter, N., Firth, C. D., Johnstone, E. C., Owens, D. E., and Roberts, G. W., Schizophrenia as an anomaly of development of cerebral asymmetry, Arch. Gen. Psychiatry, Vol. 46, pp. 1145–1150, 1989.

[75] Maes, F., Van Leemput, K., DeLisi, L. E., Vandermeulen, D., and Suetens, P., Quantification of cerebral grey and white matter asymmetry from MRI, In: Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI'99, Lecture Notes in Computer Science, Vol. 1679, Springer, Berlin, pp. 348–357, 1999.

[76] Marais, P., Guillemaud, R., Sakuma, M., Zisserman, A., and Brady, M., Visualising cerebral asymmetry, In: Visualization in Biomedical Computing, Vol. 1131 of Lecture Notes in Computer Science, Höhne, K. H. and Kikinis, R. eds., Homburg, Germany, Springer, pp. 411–416, 1999.

[77] Liu, Y., Collins, R. T., Rothfus, W. E., Automatic Bilateral symmetry (midsagittal) plane extraction from pathological 3D neuroradiological images, In: Medical Imaging 1998: Image Processing, Vol. 3338 of Proc. SPIE, Hanson, K. M. ed., San Diego, CA, USA, pp. 1528–1539, 1998.

[78] Prima, S., Thirion, J.-P., Subsol, G., and Roberts, N., Automatic analysis of normal brain dissymmetry of males and females in MR images, In: Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), Vol. 1496 of Lecture Notes in Computer Science, Wells, W. M., Colchester, A., and Delp, S., eds., Cambridge, MA, USA, Springer, pp. 770–779, 1998.

[79] Ashburner, J. and Friston, K., Voxel-based morphometry—The methods, NeuroImage, Vol. 11, No. 6, pp. 805–821, 2000.

[80] Pohl, K. M., Wells, W. M., III, Guimond, A., Kasai, K., Shenton, M. E., Kikinis, R., Grimson, W. E. L., and Warfield, S. K., Incorporating non-rigid registration into expectation maximization algorithm to segment MR images, In: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part I, Springer-Verlag, Berlin, pp. 564–571, 2002.

[81] Wyatt, P. P. and Noble, J. A., MAP MRF joint segmentation and registration of medical images, Med. Image Anal., Vol. 7, No. 4, pp. 539–552, 2003.

[82] D'Agostino, E., Maes, F., Vandermeulen, D., and Suetens, P., An information theoretic approach for non-rigid image registration using voxel class probabilities, In: Proceedings of the Second International Workshop on Biomedical Image Registration, WBIR 2003, Lecture Notes in Computer Science, Vol. 2717, pp. 122–131, Springer, 2003.

[83] Holmes, C., Hoge, R., Collins, L., Woods, R., Toga, A., and Evans, A., Enhancement of MR images using registration for signal averaging, J. Comput Tomography, Vol. 22, 1998.

[84] Kochunov, P., Lancaster, J., Thompson, P., Toga, A., Brewer, P., Hardies, J., and Fox, P., An optimized individual target brain in the Talairach coordinate system, NeuroImage, Vol. 17, 2002.

[85] Vandermeulen, D., Descombes, X., Suetens, P., and Marchal, G., Unsupervised regularized classification of multi-spectral MRI, Technical Report KUL/ESAT/MI2/9608, Katholieke Universiteit Leuven, Feb. 1996.

Chapter 2

Supervised Texture Classification for Intravascular Tissue Characterization

Oriol Pujol[1] *and Petia Radeva*[1]

2.1 Introduction

Vascular disease, stroke, and arterial dissection or rupture of coronary arteries are considered some of the main causes of mortality in present days. The behavior of the atherosclerotic lesions depends not only on the degree of lumen narrowing but also on the histological composition that causes that narrowing. Therefore, tissue characterization is a fundamental tool for studying and diagnosing the pathologies and lesions associated to the vascular tree.

Although important, tissue characterization is an arduous task that requires manual identification by specialists of the tissues and proper tissue visualization. Intravascular ultrasound (IVUS) imaging is a well suited visualization technique for such task as it provides a cross-sectional cut of the coronary vessel, unveiling its histological properties and tissue organization.

IVUS is a widespread technique accepted in clinical practice to fill up the lack of information provided by classical coronary angiography on vessel morphology. It has a prominent role evaluating the artery lesion after a interventional coronary procedure such as balloon dilation of the vessel, stent implantation, laser angioplasty, or atherectomy.

IVUS displays the morphology and histological properties of a cross section of a vessel [1]. Figure 2.1 shows a good example of different tissues in an

[1]Computer Vision Center, Universitat Autònoma de Barcelona, Campus UAB, Edifici O, 08193 Bellaterra (Barcelona), Spain

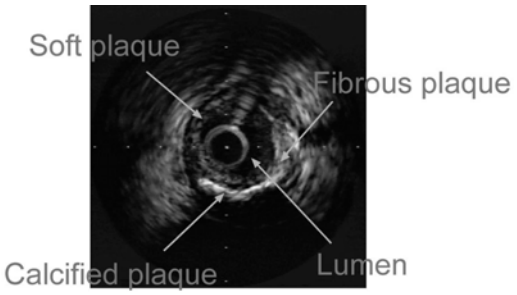


Figure 2.1: Typical IVUS image presenting different kind of tissues.

IVUS image. It is generally accepted that the different kind of plaque tissues distinguishable in IVUS images is threefold: *Calcium formation* is characterized by a very high echoreflectivity and absorbtion of the emitted pulse from the transducer. This behavior produces a deep shadowing effect behind calcium plaques. In the figure, calcium formation can be seen at three o'clock and from five to seven o'clock. *Fibrous plaque* has medium echoreflectivity resembling that of the adventitia. This tissue has a good transmission coefficient allowing the pulse to travel through the tissue, and therefore, providing a wider range of visualization. This kind of tissue can be observed from three o'clock to five o'clock. *Soft plaque* or *Fibro-Fatty plaque* is the less echoreflective of the three kind of tissues. It also has good transmission coefficient allowing to see what is behind this kind of plaque. Observing the figure, a soft plaque configuration is displayed from seven o'clock to three o'clock.

Because of time consumption and subjectivity of the classification depending on the specialist, there is a crescent interest of the medical community in developing automatic tissue characterization procedures. This is accentuated because the procedure for tissue classification by physicians implies the manual analysis of IVUS images.

The problem of automatic tissue characterization has been widely studied in different medical fields. The unreliability of gray-level only methods to achieve good discrimination among the different kind of tissues forces us to use more complex measures, usually based on texture analysis. Texture analysis has played a prominent role in computer vision to solve tissue characterization problems in medical imaging [2–9].

Several researching groups have reported different approximations to characterize the tissue of IVUS image.

Vandenberg in [10] base their contribution on reducing the noise of the image to have a clear representation of the tissue. The noise reduction is achieved by averaging sets of images when the least variance in diameter of the IVUS occurs. At the end, a fuzzy logic based expert is set to discriminate among the tissues.

Nailon and McLaughlin devote several efforts to IVUS tissue characterization. In [11] they use classic Haralick texture statistics to discriminate among tissues. In [12] the authors propose the use of co-occurrence matrices texture analysis and fractal texture analysis to characterize intravascular tissue. Thirteen features plus fractal dimension derived from Brownian motion are used for this task. The conclusion shows that fractal dimension is unable to discriminate between calcium and fibrous plaque but helps in fibrous versus lipidic plaque. On the other hand, co-occurrence matrices are well suited for the overall classification. In [13], it is stated that the discriminative power of fractal dimension is poor when trying to separate fibrotic tissue, lipidic tissue, and foam cells. The method used is based on fractal dimension estimation techniques (box-counting, brownian motion, and frequency domain).

Spencer in [14] center their work on spectral analysis. Different features are compared: mean power, maximum power, spectral slope, and 0 Hz interception. The work concludes with the 0 Hz spectral slope as the most discriminative feature.

Dixon in [15] use co-occurrence matrices and discriminant analysis to evaluate the different kind of tissues in IVUS images.

Ahmed and Leyman in [16] use a radial transform and correlation for pattern matching. The features used are higher order statistics such as kurtosis, skewness, and up to four order cumulants. The results provided appear to have fairly good visual recognition rate.

The work of de Korte and van der Steen [17] opens a new proposal based on assessing the local strain of the atherosclerotic vessel wall to identify different plaque components. This very promising technique, called *elastography*, is based on estimating the radial strain by performing cross-correlation analysis on pairs of IVUS at a certain intracoronary pressure.

Probably, one of the most interesting work in this field is the one provided by Zhang and Sonka in [18]. This work is much more complex trying to evaluate the full morphology of the vessel. Detecting the plaque and adventitia borders and characterizing the different kind of tissues, the tissue discrimination is done using a combination of well-known techniques previously reported in the

literature, co-occurrence matrices and fractal dimension from brownian motion, and adding two more strategies to the amalgam of features: run-length measures and radial profile. The experiments assess the accuracy of the method quantitatively.

Most of the literature found in the tissue characterization matters use texture features, co-occurrence matrices being the most popular of all feature extractors. Further work has been done trying to use other kind of texture feature extractors and IVUS images, and although not specifically centered on tissue characterization, the usage of different texture features in plaque border assessment is reported, which can be easily extrapolated to tissue characterization. In [19], derivatives of Gaussian; wavelets, co-occurrence matrices, Gabor filters, and accumulation local moments are evaluated and used to classify blood from plaque. The work highlights the discriminative power of co-occurrence matrices derivatives of Gaussian and accumulation local moments. Other works such as [20] provide some hints on how to achieve a fast framework based on local binary patterns and fast high-performance classifiers. This last line of investigation overcomes one of the most significant drawbacks of the texture based tissue characterization systems, the speed. Texture descriptors are inherently slow to be computed. With the proposal of the feature extractor based on local binary patterns a good discriminative power is ensured as well as a fast technique for tissue characterization.

Whatever method we use in the tissue characterization task, we follow an underlying main methodology. First, we need to extract some features describing the tissue variations. This first step is critical since the features chosen have to be able to describe each kind of tissue in a unique way so that it cannot be confused with another one. In this category of feature extraction we should consider the *co-occurrence matrix measures*, *local binary patterns*, etc. The second step is the classification of the extracted features. Depending on the complexity of the feature data some methods will fit better than others. In most cases, high-dimensional spaces are generated, so we should consider the use of dimensionality reduction methods such as *principal component analysis* or *Fisher linear discriminant analysis*. Either a dimensionality reduction process is needed or not, this step will require a classification procedure. This procedure can be supervised, if we provide samples of each tissue to be classified so that the system "knows" *a priori* what the tissues are, or unsupervised, if we allow the system to try to find which are the different kind of tissues by itself. In

this category, we can find *clustering* methods for unsupervised classification and, for supervised classification, methods like *maximum likelihood*, *nearest neighbors*, etc.

The following sections are devoted to describe the following: First, the most significative texture methods used in the literature. Secondly, some of the most successful classification methods applied to IVUS characterization are reviewed. Third, we describe the result of using such techniques for tissue characterization and conclude about the optimal feature space to describe tissue plaque and the best classifier to discriminate it.

2.2 Feature Spaces

Gray-level thresholding is not enough for robust tissue characterization. Therefore, it is generally approached as a texture discrimination problem. This line of work is a classical extension of previous works on biological characterization, which also relies on texture features as has been mentioned in the former section. The co-occurrence matrix is the most favored and well known of the texture feature extraction methods due to its discriminative power in this particular problem but it is not the only one nor the fastest method available. In this section, we make a review of different texture methods that can be applied to the problem in particular, from the *co-occurrence matrix measures* method to the most recent texture feature extractor, *local binary patterns*.

To illustrate the texture feature extraction process we have selected a set of techniques basing our criterion of selection on the most widespread methods for tissue characterization and the most discriminative feature extractors reported in the literature [21].

Basically, the different methods of feature extraction emphasize on different fundamental properties of the texture such as scale, statistics, or structure. In this way, under the nonelemental statistics property we can find two well-known techniques, co-occurrence methods [22] and higher order statistics represented by moments [23]. Under the label of scale property we should mention methods such as derivatives of Gaussian [24], Gabor filters [25], or wavelet techniques [26]. Regarding structure-related measures there are methods such as fractal dimension [27] and local binary patterns [28].

To introduce the texture feature extraction methods we divide them into two groups: The first group, that forms the *statistic-related methods*, is comprised of co-occurrence matrix measures, accumulation local moments, fractal dimension, and local binary patterns. All these methods are somehow related to statistics. Co-occurrence matrix measures are second-order measures associated to the probability density function estimation provided by the co-occurrence matrix. Accumulation local moments are directly related to statistics. Fractal dimension is an approximation of the roughness of a texture. Local binary patterns provides a measure of the local inhomogeneity based on an "averaging" process. The second group, that forms the *analytic kernel-based extraction techniques*, comprises Gabor bank of filters, derivatives of Gaussian filters, and wavelet decomposition. The last three methods are derived from analytic functions and sampled to form a set of filters, each focused on the extraction of a certain feature.

2.2.1 Statistic-Related Methods

2.2.1.1 Co-occurrence Matrix Approach

In 1962 Julesz [29] showed the importance of texture segregation using second-order statistics. Since then, different tools have been used to exploit this issue. The gray-level co-occurrence matrix is a well-known statistical tool for extracting second-order texture information from images [22]. In the co-occurrence method, the relative frequencies of gray-level pairs of pixels at certain relative displacement are computed and sorted in a matrix, the *co-occurrence matrix* **P**. The co-occurrence matrix can be thought of as an estimate of the joint probability density function of gray-level pairs in an image. For G gray levels in the image, **P** will be of size $G \times G$. If G is large, the number of pixel pairs contributing to each element, $p_{i,j}$ in **P** is low, and the statistical significance poor. On the other hand, if the number of gray levels is low, much of the texture information may be lost in the image quantization. The element values in the matrix, when normalized, are bounded by $[0, 1]$, and the sum of all element values is equal to 1.

$$P(i, j, D, \theta) = P(I(l, m) = i \quad \text{and} \quad I(l + D\cos(\theta), m + D\sin(\theta)) = j$$

where $I(l, m)$ is the image at pixel (l, m), D is the distance between pixels,

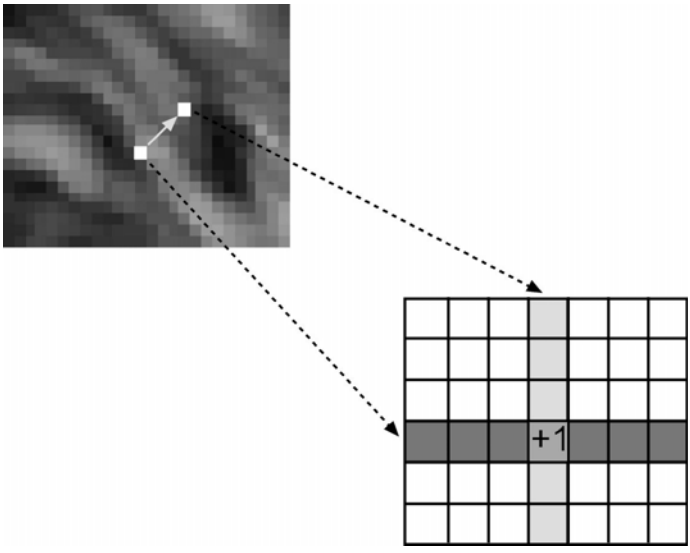


Figure 2.2: Co-occurrence matrix explanation diagram (see text).

and θ is the angle. It has been proved by other researchers [21, 30] that the nearest neighbor pairs at distance D at orientations $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ are the minimum set needed to describe the texture second-order statistic measures. Figure 2.2 illustrates the method providing a graphical explanation. The main idea is to create a "histogram" of the occurrences of having two pixels of certain gray levels at a determined distance with a fixed angle. Practically, we add one to the cell of the matrix pointed by the gray levels of two pixels (one pixel gray level gives the file and the other the column of the matrix) that fulfill the requirement of being at a certain predefined distance and angle.

Once the matrix is computed several characterizing measures are extracted. Many of these features are derived by weighting each of the matrix element values and then summing these weighted values to form the feature value. The weight applied to each element is based on a feature-weighing function, so by varying this function, different texture information can be extracted from the matrix. We present here some of the most important measures that characterize the co-occurrence matrices: energy, entropy, inverse difference moment, shade, inertia, and promenance [30]. Let us introduce some notation for the definition of the features:

$P(i, j)$ is the (i, j)th element of a normalized co-occurrence matrix

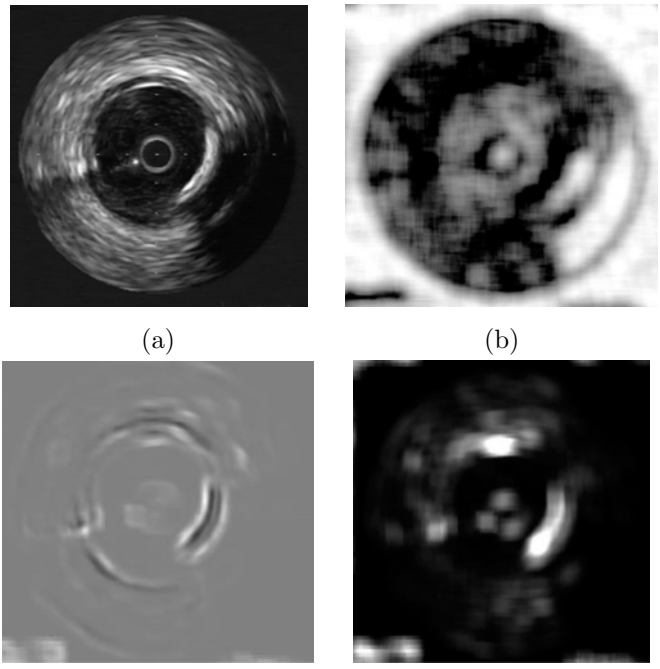
$$
\begin{aligned}
P_x(i) &= \sum_j P(i, j) \\
P_y(j) &= \sum_i P(i, j) \\
\mu_x &= \sum_i i \sum_j P(i, j) = \sum_i i P_x(i) = E\{i\} \\
\mu_y &= \sum_j j \sum_i P(i, j) = \sum_j j P_y(j) = E\{j\}
\end{aligned}
$$

With the above notation, the features can be written as follows:

$$
\begin{aligned}
\text{Energy} &= \sum_{i,j} P(i, j)^2 \\
\text{Entropy} &= -\sum_{i,j} P(i, j) \log P(i, j) \\
\text{Inverse difference moment} &= \sum_{i,j} \frac{1}{1 + (i - j)^2} P(i, j) \\
\text{Shade} &= \sum_{i,j} (i + j - \mu_x - \mu_y)^3 P(i, j) \\
\text{Inertia} &= \sum_{i,j} (i - j)^2 P(i, j) \\
\text{Promenance} &= \sum_{i,j} (i + j - \mu_x - \mu_y)^4 P(i, j)
\end{aligned}
$$

Hence, we create a *feature vector* for each of the pixels by assigning each feature measure to a component of the feature vector. Given that we have four different orientations and the six measures for each orientation, the feature vector is a 24-dimensional vector for each pixel and for each distance. Since we have used two distances $D = 2$ and $D = 3$, the final vector is a 48-dimensional vector.

Figure 2.3 shows responses for different measures on the co-occurrence matrices. Although a straightforward interpretation of the feature extraction response is not easy, some deduction can be made by observing the figures. Figure 2.3(b) shows shade measure; as its name indicates it is related to the shadowed areas in the image, and thus, localizing the shadowing behind the calcium plaque. Figure 2.3(c) shows inverse different moment response, this measure seems to be related to the first derivative of the image, enhancing contours. Figure 2.3(d) depicts the output for the inertia measure, which seems to have some relationship with local homogeneity of the image.



(a) (b)

(c) (d)

Figure 2.3: Response of an IVUS image to different measures of the co-occurrence matrix. (a) Original image, (b) measure shade response, (c) inverse different moment, and (d) inertia.

2.2.1.2 Accumulation Local Moments

Geometric moments have been used effectively for texture segmentation in many different application domains [23]. In addition, other kind of moments have been proposed: Zernique moments, Legendre moments, etc. By definition, any set of parameters obtained by projecting an image onto a two-dimensional polynomial basis is called moments. Then, since different sets of polynomials up to the same order define the same subspace, any complete set of moments up to given order can be obtained from any other set of moments up to the same order. The computation of some of these sets of moments leads to very long processing times, so in this section a particular fast computed moment set has been chosen. This set of moments is known as the accumulation local moments. Two kind of accumulation local moments can be computed, direct accumulation and reverse accumulation. Since direct accumulation is more sensitive to round

off errors and small perturbations in the input data [31], the reverse accumulation moments are recommendable.

The reverse accumulation moment of order $(k-1, l-1)$ of matrix $\mathbf{I}_{ab}$ is the value of $\mathbf{I}_{ab}[1, 1]$ after bottom-up accumulating its column k times (i.e., after applying k times the assignment $\mathbf{I}_{ab}[a-i, j] \leftarrow \mathbf{I}_{ab}[a-i, j] + \mathbf{I}_{ab}[a-i+1, j]$, for $i = 0$ to $a-1$, and for $j = 1$ to b), and accumulating the resulting first row from right to left l times (i.e., after applying l times the assignment $\mathbf{I}_{ab}[1, b-j] \leftarrow \mathbf{I}_{ab}[1, b-j] + \mathbf{I}_{ab}[1, b-j+1]$, for $j = 1$ to $b-1$). The reverse accumulation moment matrix is defined so that $\mathbf{R}_{mn}[k.l]$ is the reverse accumulation moment of order $(k-1, l-1)$.

Consider the matrix in the following example:

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 4 & 2 & 3 \end{pmatrix}$$

According to the definition, its reverse accumulation moment of order (1,2) requires two column accumulations,

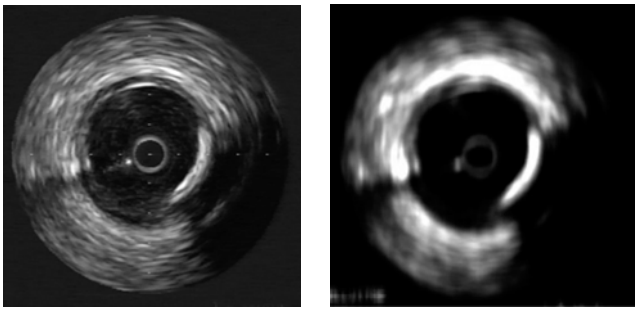
$$\begin{pmatrix} 5 & 4 & 6 \\ 5 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 14 & 9 & 13 \\ 9 & 5 & 7 \\ 4 & 2 & 3 \end{pmatrix}$$

and three right to left accumulations of the first row:

$$\begin{pmatrix} 36 & 22 & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 71 & 35 & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 119 & 48 & 13 \end{pmatrix}$$

Then it is said that the reverse accumulation moment of order (1,2) of the former matrix is 119.

The set of moments alone is not sufficient to obtain good texture features in certain images. Some iso-second order texture pairs, which are preattentively discriminable by humans, would have the same average energy over finite regions. However, their distribution would be different for the different textures. One solution suggested by Caelli is to introduce a nonlinear transducer that maps moments to texture features [32]. Several functions have been proposed in the literature: logistic, sigmoidal, power function, or absolute deviation of feature vectors from the mean [23]. The function we have chosen is the hyperbolic tangent function, which is logistic in shape. Using the accumulation moments image I_m and a nonlinear operator $|\tanh(\sigma(I_m - \bar{I}_m)|$ an "average" is performed



(a) (b)

Figure 2.4: Accumulation local moments response. (a) Original image. (b) Accumulation local moment of order (3,1).

throughout the region of interest. The parameter σ controls the shape of the logistic function. Therefore each textural feature will be the result of the application of the nonlinear operator to the computed moments. If $n = k \cdot l$ moments are computed over the image, then the dimension of the feature vector will be n. Hence, a n-dimensional point is associated with each pixel of the image.

Figure 2.4 shows the response of moment (3,1) on an IVUS image. In this figure, the response seems to have a smoothing and enhancing effect, clearly resembling diffusion techniques.

2.2.1.3 Fractal Analysis

Another classic tool for texture description is the fractal analysis [13, 33], characterized by the fractal dimension. We talk roughly about fractal structures when a geometric shape can be subdivided in parts, each of which are approximately a reduced copy of the whole (this property is also referred as self-similarity). The introduction of fractals by Mandelbrot [27] allowed a characterization of complex structures that could not be described by a single measure using Euclidean geometry. This measure is the *fractal dimension*, which is related to the degree of irregularity of the surface texture.

The fractal structures can be divided into two subclasses: the deterministic fractals and the random fractals. Deterministic fractals are strictly self-similar, that is, they appear identical over a range of magnification scales. On the other hand, random fractals are statistical self-similar. The similarity between two scales of the fractal is ruled by a statistical relationship.

The fractal dimension represents the disorder of an object. The higher the dimension, the more complex the object is. Contrary to the Euclidian dimension, the fractal dimension is not constrained to integer dimensions.

The concept of fractals can be easily extrapolated to image analysis if we consider the image as a three-dimensional surface in which the height at each point is given by the gray value of the pixel.

Different approaches have been proposed to compute the fractal dimension of an object. Herein we consider only three classical approaches: box-counting, Brownian motion, and Fourier analysis.

Box-Counting. The box-counting method is an approximation to the fractal dimension as it is conceptually related to self-similarity.

In this method the object to be evaluated is placed on a square mesh of various sizes, r. The number of mesh boxes, N, that contain any part of the fractal structure are counted.

It has been proved that in a self-similar structures there is a relationship between the reduction factor r and the number of divisions N into which the structure can be divided:

$$Nr^D = 1$$

where D is the self-similarity dimension. Therefore, the fractal dimension can be easily written as

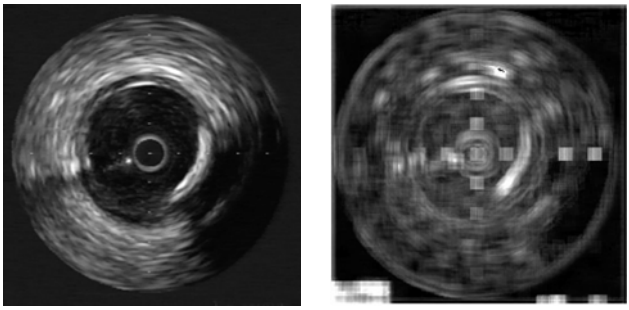
$$D = \frac{\log N}{\log 1/r}$$

This process is done at various scales by altering the square size r. Therefore, the box-counting dimension is the slope of the regression line that better approximates the data on the plot produced by $\log N \times \log 1/r$.

Fractal Dimension from Brownian Motion. The fractal dimension is found by considering the absolute intensity difference of pixel pairs, $I(p_1) - I(p_2)$, at different scales. It can be shown that for a fractal Brownian surface the following relationship must be satisfied:

$$E(|I(p_1) - I(p_2)|)\alpha(\sqrt{(x_2 - x_1) + (y_2 - y_1)})^H$$

where E is the mean and H the Hurst coefficient. The fractal dimension is related to H in the following way: $D = 3 - H$. In the same way than the former



(a) (b)

Figure 2.5: Fractal dimension from box-counting response. (a) Original image. (b) Fractal dimension response with neighborhoods of 10×10.

method for calculating the fractal dimension the mean difference of intensities is calculated for different scales (each scale given by the euclidian distance between two pixels), and the slope of the regression line between $\log E(|I(p_1) - I(p_2)|)$ and $\sqrt{(x2-x1)+(y2-y1)}$ gives the Hurst parameter.

Triangular Prism Surface Area Method. The triangular prism surface area (TPSA) algorithm considers an approximation of the "area" of the fractal structure using triangular prisms. If a rectangular neighborhood is defined by its vertices A, B, C, and D, the area of this neighborhood is calculated by tessellating the surface with four triangles defined for each consecutive vertex and the center of the neighborhood.

The area of all triangles for every central pixel is summed up to the entire area for different scales. The double logarithmic Richardson–Mandelbrot plot should again yield a linear line whose slope is used to determine the TPSA dimension. Figure 2.5 shows the fractal dimension value of each pixel of an IVUS image considering the fractal dimension of a neighborhood around the pixel. The size of the neighborhood is 10×10. The response of this technique seems to take into account the border information of the structures in the image.

2.2.1.4 Local Binary Patterns

Local binary patterns [28] are a feature extraction operator used for detecting "uniform" local binary patterns at circular neighborhoods of any quantization of

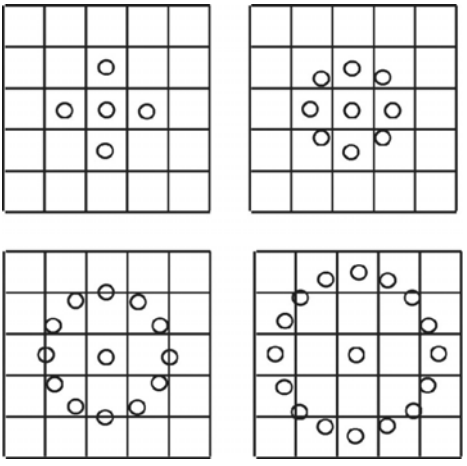


Figure 2.6: Typical neighbors: (Top left) $P = 4$, $R = 1.0$; (top right) $P = 8$, $R = 1.0$; (bottom left) $P = 12$, $R = 1.5$; (bottom right) $P = 16$, $R = 2.0$.

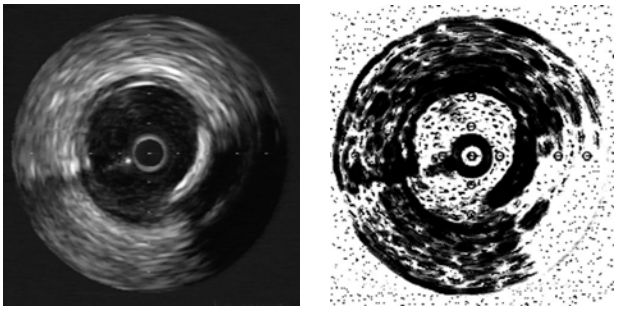
the angular space and at any spatial resolution. The operator is derived based on a circularly symmetric neighbor set of P members on a circle of radius R. It is denoted by $LBP^{riu2}_{P,R}$. Parameter P controls the quantization of the angular space, and R determines the spatial resolution of the operator. Figure 2.6 shows typical neighborhood sets. To achieve gray-scale invariance, the gray value of the center pixel (g_c) is subtracted from the gray values of the circularly symmetric neighborhood g_p ($p = 0, 1, \ldots, P - 1$) and assigned a value of 1 if the difference is positive and 0 if negative.

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

By assigning a binomial factor 2^p for each value obtained, we transform the neighborhood into a single value. This value is the $LBP_{R,P}$:

$$LBP_{R,P} = \sum_{p=0}^{P} s(g_p - g_c) \cdot 2^p$$

To achieve rotation invariance the pattern set is rotated as many times as necessary to achieve a maximal number of the most significant bits, starting always from the same pixel. The last stage of the operator consists on keeping the information of "uniform" patterns while filtering the rest. This is achieved using a transition count function U. U is a function that counts the number of transitions



(a) (b)

Figure 2.7: Local binary pattern response. (a) Original image. (b) Local binary pattern output with parameters $R = 3$, $P = 24$.

$0/1$, $1/0$ while we move over the neighborhood:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$$

Therefore,

$$LBP^{riu2}_{P,R} = \begin{cases} LBP^{ri}_{P,R} & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise.} \end{cases}$$

Figure 2.7 shows an example of an IVUS image filtered using a uniform rotation invariant local binary pattern with values $P = 24$, $R = 3$. The feature extraction image displayed in the figure looks like a discrete response focussed on the structure shape and homogeneity.

2.2.2 Analytic Kernel-Based Methods

2.2.2.1 Derivatives of Gaussian

In order to handle image structures at different scales in a consistent manner, a linear *scale-space representation* is proposed in [24, 34]. The basic idea is to embed the original signal into an one-parameter family of gradually smoothed signals, in which fine scale details are successively suppressed. It can be shown that the Gaussian kernel and its derivatives are one of the possible smoothing kernels for such scale-space. The Gaussian; kernel is well-suited for defining a space-scale because of its linearity and spatial shift invariance, and the notion

that structures at coarse scales should be related to structures at finer scales in a well-behaved manner (new structures are not created by the smoothing method). Scale-space representation is a special type of multiscale representation that comprises a continuous scale parameter and preserves the same spatial sampling at all scales. Formally, the linear-space representation of a continuous signal is constructed as follows. Let $f : \Re^N \rightarrow \Re$ represent any given signal. Then, the scale-space representation $L : \Re^N \times R_+ \rightarrow \Re$ is defined by $L(\cdot; 0) = f$ so that

$$L(\cdot; t) = g(\cdot; t) * f$$

where $t \in \Re_+$ is the scale parameter, and $g : \Re^N x R_+ \{0\} \rightarrow \Re$ is the Gaussian kernel. In arbitrary dimensions, it is written as:

$$g(x; t) = \frac{1}{(2\pi t)^{N/2}} \, e^{-x^T x/(2t)} = \frac{1}{(2\pi t)^{N/2}} \, e^{-\sum_{i=1}^{N} x_i^2/(2t)} \quad x \in Re^N, x_i \in \Re$$

The square root of the scale parameter, $\sigma = \sqrt{(t)}$, is the standard deviation of the kernel g and is a natural measure of spatial scale in the smoothed signal at scale t. From this scale-space representation, multiscale spatial derivatives can be defined by

$$L_{x^n}(\cdot; t) = \partial_{x^n} L(\cdot; t) = g_{x^n}(\cdot; t) * f,$$

where g_{x^n} denotes a derivative of some order n.

The main idea behind the construction of this scale-space representation is that the fine scale information should be suppressed with increasing values of the scale parameter. Intuitively, when convolving a signal by a Gaussian kernel with standard deviation $\sigma = \sqrt{t}$, the effect of this operation is to suppress most of the structures in the signal with a characteristic length less than σ. Different directional derivatives can be used to extract different kind of structural features at different scales. It is shown in the literature [35] that a possible complete set of directional derivatives up to the third order are $\partial^n = [\partial_0, \partial_{90}, \partial_0^2, \partial_{60}^2, \partial_{120}^2, \partial_0^3, \partial_{45}^3, \partial_{90}^3, \partial_{135}^3]$. So our feature vector will consist on the directional derivatives, including the zero derivative, for each of the n scales desired:

$$F = \{\{\partial^n, G^n\}, n \in \Re\}$$

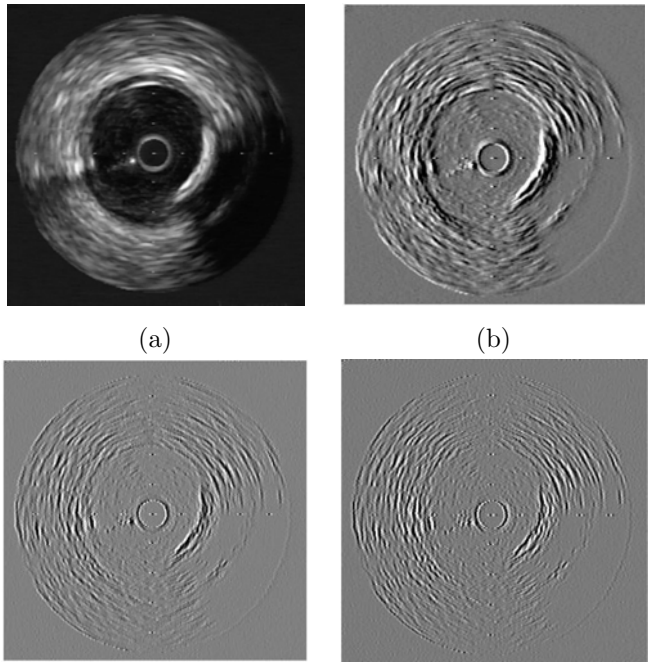


Figure 2.8: Derivative of Gaussian responses for $\sigma = 2$. (a) Original image; (b) first derivative of Gaussian response; (c) second derivative of Gaussian response; (d) third derivative of Gaussian response.

Figure 2.8 shows some of the responses for the derivative of Gaussian bank of filters for $\sigma = 2$. Figures 2.8(b), 2.8(c), and 2.8(d) display the first, second, and third derivatives of Gaussian, respectively.

2.2.2.2 Wavelets

Wavelets come to light as a tool to study nonstationary problems [36]. Wavelets perform a decomposition of a function as a sum of local bases with finite support and localized at different scales. Wavelets are characterized for being bounded functions with zero average. This implies that the shapes of these functions are waves restricted in time. Their time-frequency limitation yields a good location. So a wavelet ψ is a function of zero average:

$$\int_{-\infty}^{+\infty} \psi(t)\, dt = 0$$

which is dilated with a scale parameter s and translated by u:

$$\varphi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)$$

The wavelet transform of f at scale s and position u is computed by correlating f with a wavelet atom:

$$W_f(u,s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{(s)}}\psi^*\left(\frac{t-u}{s}\right)dt \tag{2.1}$$

The continuous wavelet transform $W_f(u,s)$ is a two-dimensional representation of a one-dimensional signal f. This indicates the existence of some redundancy that can be reduced and even removed by subsampling the parameters of these transforms. Completely eliminating the redundancy is equivalent to building a basis of the signal space.

The decomposition of a signal gives a series of coefficients representing the signal in terms of the base from a mother wavelet, that is, the projection of the signal on the space formed by the base functions.

The continuous wavelet transform has two major drawbacks: the first, stated formerly, is redundancy and the second, impossibility to calculate it unless a discrete version is used. A way to discretize the dilation parameter is $a = a_0^m$, $m \in Z$, $a_0 \neq 1$ constant. Thus, we get a series of wavelets ψ_m of width, a_0^m. Usually, we take $a_0 > 1$, although it is not important because m can be positive or negative. Often, a value of $a_0 = 2$ is taken. For $m = 0$, we make s to be the only integer multiples of a new constant s_0. This constant is chosen in such a way that the translations of the mother wavelet, $\psi(t - ns_0)$, are as close as possible in order to cover the whole real line. Then, the election of s level is as follows:

$$\psi_m, n(t) = a_0^{-m/2}\psi\left(\frac{t - ns_0a^m}{a_0^m}\right) = a_0^{-m/2}\psi\left(a_0^{-m}t - ns_0\right)$$

that covers the entire real axis as well as the translations $\psi(t - ns_0)$ does. Summarizing, the discrete wavelet transform consists of two discretizations in the transformation Eq. (2.1),

$$a = a_0^m, \qquad b = nb_0a_0^m, \qquad m, n \in Z, \quad a_0 > 1, b_0 > 0$$

The multiresolution analysis (MRA) tries to build orthonormal bases for a dyadic grid, where $a_0 = 2, b_0 = 1$, which besides have a compact support region. Finally, we can imagine the coefficients $d_{m,n}$ of the discrete wavelet transform as the sampling of the convolution of signal $f(t)$ with different filters $\psi_m(-t)$,

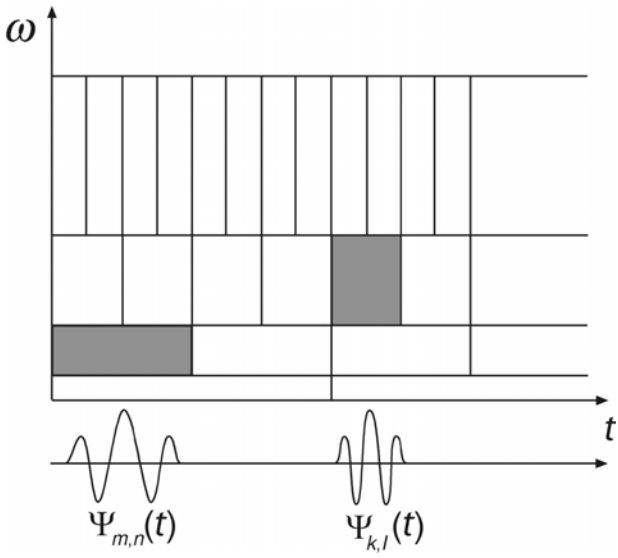


Figure 2.9: Scale-frequency domain of wavelets.

where $\psi_m(t) = a_0^{-m/2}\psi(a^{-m}t)$

$$y_m(t) = \int f(s)\psi_m(s-t)\,ds \qquad d_{m,n} = y_m\left(na_0^m\right)$$

Figure 2.9 shows the dual effect of shrinking of the mother wavelet as the frequency increases, and the translation value decreasing as the frequency increases. The mother wavelet keeps its shape but if high-frequency analysis is desired the spatial support of the wavelet has to decrease. On the other hand, if the whole real line has to be covered by translations of the mother wavelet, as the spatial support of the wavelet decreases, the number of translations needed to cover the real line increases. Unlike Fourier transform, where translations of analysis are at the same distance for all the frequencies.

The choice of a representation of the wavelet transform leads us to define the concept of a *frame*. A frame is a complete set of functions that, though able to span $L^2(\Re)$, is not a base because it lacks the property of linear independence. MRA proposed in [26] is another representation in which the signal is decomposed in an approximation at a certain level L with L detail terms of higher resolutions. The representation is an orthonormal decomposition instead of a redundant frame, and therefore, the number of samples that defines a signal is the same as that the number of coefficients of their transform. A MRA consists

of a sequence of function subspaces of successive approximation. Let P_j be an operator defined as the orthonormal projection of functions of L^2 over the space V_j. The projection of a function f over V_j is a new function that can be expressed as a linear combination of the functions that form the orthonormal base of V_j. Coefficients of the combination of each base function is the scalar product of f with the base functions:

$$P_j f = \sum_{n \in Z} \langle f, \phi_{j,n} \rangle \phi_{j,n}$$

where

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t) g(t)\, dt$$

Earlier we have pointed out the nesting condition of the V_j spaces, $V_j \subset V_{j-1}$. Now, if $f \in V_{j-1}$ then $f \in V_j$ or f is orthonormal to all the V_j functions, that is, we divide V_{j-1} in two disjoint parts: V_j and other space W_j, such that if $f \in V_j, g \in W_j$, $f \perp g$; W_j is the orthonormal complement of V_j in V_{j-1}:

$$V_{j-1} = V_j \oplus W_j$$

where symbol $\oplus$ measures the addition of orthonormal spaces. Applying the former equation and the completeness condition, then

$$\cdots \oplus W_{j-2} \oplus W_{j-1} \oplus W_j \oplus W_{j+1} \oplus \cdots = \bigoplus_{j \in Z} W_j = L^2$$

So, we can write

$$P_{j-1} f = P_j f + \sum_{n \in Z} \langle f, \psi_{j,n} \rangle \psi_{j,n}$$

From these equations some conclusions can be extracted. First, the projection of a signal f in a space V_j gives a new signal $P_j f$, an approximation of the initial signal. Secondly, we have a hierarchy of spaces, then $P_{j-1} f$ will be a better *approximation* (more reliable) than $P_j f$. Since V_{j-1} can be divided in two subspaces V_j and W_j, if V_j is an approximation space then W_j, which is the complementary orthonormal space, it is the *detail* space. The less the j, the finer the details.

$$\begin{aligned} V_j &= V_{j+1} \oplus W_{j+1} = V_{j+2} \oplus W_{j+2} = \cdots \\ &= V_L \oplus W_L \oplus W_{L-1} \oplus \cdots \oplus W_{j+1} \end{aligned}$$

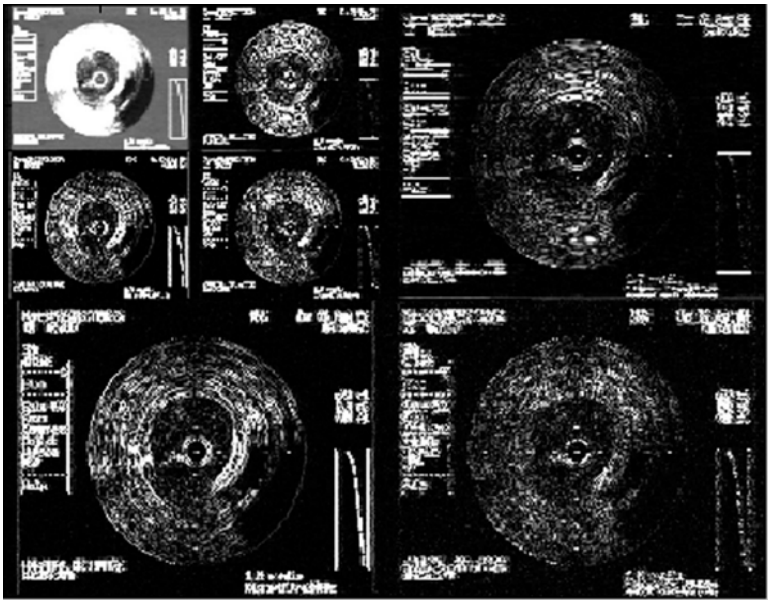


Figure 2.10: Wavelets multiresolution decomposition.

This can be viewed as a decomposition tree (see Fig. 2.10). At the top-left side of the image the approximation can be seen, and surrounding it the successive details. The further the detail is located the finer the information provided. So, the details at the bottom and at the right side of the image have information about the finer details and the smallest structures of the image decomposed. Therefore, we have a feature vector composed by the different detail approaches and the approximation for each of the pixels.

2.2.2.3 Gabor Filters

Gabor filters represent another multiresolution technique that relies on scale and direction of the contours [25,37]. The Gabor filter consists of a two-dimensional sinusoidal plane wave of a certain orientation and frequency that is modulated in amplitude by a two-dimensional Gaussian envelope. The spatial representation of the Gabor filter is as follows:

$$h(x, y) = \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right\}\cos(2\pi u_0 x + \phi)$$

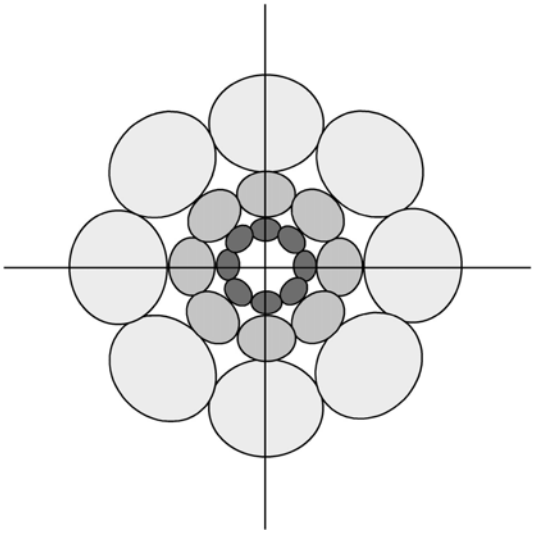


Figure 2.11: The filter set in the spatial-frequency domain.

where u_0 and ϕ are the frequency and phase of the sinusoidal plane wave along the x axis and σ_x and σ_y are the space constants of the Gaussian envelope along the x and y axis, respectively. Filters at different orientations can be created by rigid rotation of x–y coordinate system.

An interesting property of this kind of filters is their frequency and orientation-selection. This fact is better displayed in the frequency domain. Figure 2.11 shows the filter area in the frequency domain. We can observe that each of the filters has a certain domain defined by each of the leaves of the Gabor "rose." Thus, each filter responds to a certain orientation and at a certain detail level. Wider the range of orientations, smaller the space filter dimensions and smaller the details captured by the filter, as bandwidth in the frequency domain is inversely related to filter scope in the space domain. Therefore, Gabor filters provide a trade-off between localization or resolution in both the spatial and the spatial-frequential domains. As it has been mentioned, different filters emerge from rotating the x–y coordinate system. For practical approaches one can use four angles $\theta_0 = 0^\circ$, 45°, 90°, 135°. For an image array of N pixels (with N power of 2), the following values of u_0 are suggested [25, 37]:

$$1\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}, \ldots, \text{and } (N_c/4)\sqrt{2}$$

cycles per image width. Therefore, the orientations and bandwidth of such filters vary with 45° and 1 octave. These parameters are chosen because there is physiologic evidences of frequency bandwidth of simple cells in visual cortex being of about 1 octave, and Gabor filters try to mimic part of the human perceptual system.

The Gabor function is an approximation to a wavelet. However, though admissible, it does not result in an orthogonal decomposition, and therefore, a transformation based on Gabor's filters is redundant. On the other hand, Gabor filtering is designed to be nearly orthogonal, reducing the amount of overlap between filters.

Figure 2.12 shows different responses for different filters of the spectrum. Figures 2.12(a) and 2.12(b) correspond to the inner filters with reduced frequency bandwidth displayed in Fig. 2.11. It can be seen that they deliver only

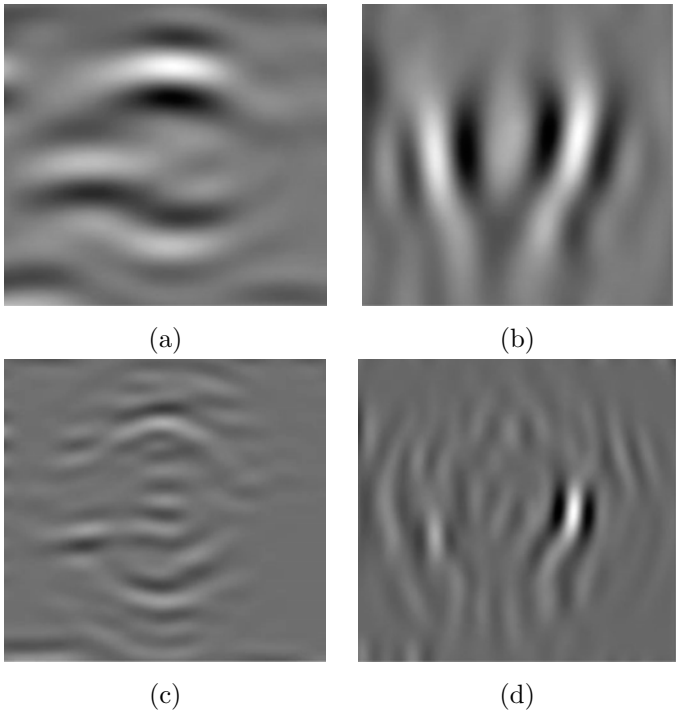


Figure 2.12: Gabor filter bank example responses. (a) Gabor vertical energy of a coarse filter response. (b) Gabor horizontal energy of a coarse filter response. (c) Gabor vertical energy of a detail filter response. (d) Gabor horizontal energy of a detail filter response.

Table 2.1: Dimensionality of the feature space provided by the texture feature extraction process

Method	Space dimension
Co-occurrence matrix measures	48
Accumulation local moments	81
Fractal Dimension	1
Local Binary Patterns	3
Derivative of Gaussian	60
Wavelets	31
Gabor's filters	20

coarse information of the structure and the borders are far from the original location. In the same way, Figs. 12(c) and 12(d) are filters located on a further ring, and therefore respond to details in the image.

It can be observed that the feature extraction process is a transformation of the original two-dimensional image domain to a feature space that probably will have different dimensions. In some cases, the feature space remains low, as in fractal dimension and local binary patterns, that with very few features try to describe the texture present in the image. However, several feature spaces require higher dimensions, such as accumulation local moments, co-occurrence matrix measures, or derivatives of Gaussian. Table 2.1 shows the dimensionality of the different spaces generated by the feature extraction process in our texture-based IVUS analysis.

The next step after the feature extraction is the classification process. As a result of the disparity of the dimensionality of the feature spaces, we have to choose a classification scheme that is able to deal with high dimensionality feature data.

2.3 Classification Process

Once completed the feature extraction process, we have a set of features disposed in feature vectors. Each feature vector is composed of all the feature measures computed at each pixel. Therefore, for each pixel we have an n-dimensional point in the feature space, where n is the number of features. This set of data is the input to the classification process. The classification process

is divided in two main categories: supervised and unsupervised learning. While supervised learning is based on a set of examples of each class that trains the classification process, the unsupervised learning is based on the geometry position of the data in the feature space and its possibility to be grouped in clusters.

In this chapter we are mainly concerned with supervised learning and classification, since we know exactly what classes we are seeking. Supervised classification techniques are usually divided in parametric and nonparametric. Parametric techniques rely on knowledge of the probability density function of each class. On the contrary, nonparametric classification, does not need the probability density function and is based on the geometrical arrangement of the points in the input space. We begin describing a nonparametric technique, *k-nearest neighbors*, that will serve as a ground truth to verify the discriminability of the different feature spaces. Since nonparametric techniques have high computational cost, we make some assumptions that lead to describe *maximum likelihood* classification techniques. However, the last techniques are very sensitive to the input space dimension. It has been shown in the former section that some feature spaces cast the two-dimensional image data to high-dimensional spaces. In order to deal with high-dimensional data, a dimensionality reduction is needed. The dimensionality reduction techniques are useful to create a meaningful set of data because the feature space is usually large in comparison to the number of samples retrieved. The most known technique for dimensionality reduction is *principal component analysis* (PCA) [38]. However, PCA is susceptible to errors depending on the arrangement of the data points in the training space, because it does not consider the different distributions of data clusters. In order to solve the deficiency of PCA in discrimination matters, *Fisher linear discriminant analysis* is introduced [38, 39]. In order to try to improve the classification rate of simple classifiers, combination of classifiers is proposed. One of the most important classification assembling process is *boosting*. The last part of this section is devoted to a particular class of boosting techniques, *Adaptative Boosting (AdaBoost)* [40, 41].

2.3.1 k-Nearest Neighbors

Voting k-nearest neighbors classification procedure is a very popular classification scheme that does not rely on any assumption concerning the structure of the underlying density function.

As any nonparametric technique, the resulting classification error is the smallest achievable error given a set of data. This is true because this technique implicitly estimates the density function of the data, and therefore, the classifier becomes the *Bayes classifier* if the density estimates converge to the true densities when an infinite number of samples are used [38].

In order to classify a test sample X, the k-nearest neighbors to the test sample are selected from the overall training data, and the number of neighbors from each class ω_i among the k selected samples is counted. The test sample is then classified to the class represented by a majority of the k-nearest neighbors. That is

$$\mathbf{k}_j = \max\{\mathbf{k}_1 \cdots \mathbf{k}_L\} \rightarrow X \in \omega_j$$
$$\mathbf{k}_1 + \cdots + \mathbf{k}_L = k$$

where $\mathbf{k}_j$ is the number of neighbors from class ω_j, $(j = 1, \ldots, L)$ among the selected neighbors. Usually, the same metric is used to measure the distance to samples of each class.

Figure 2.13 shows an example of a 5-nearest neighbors process. Sample X will be classified as member of the light gray class since there are 3-nearest neighbors of the black class while there are only 2 members of the white class.

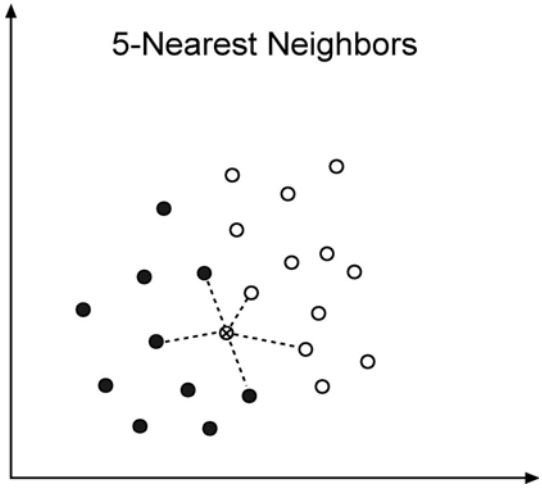


Figure 2.13: A 5-nearest neighbors example.

2.3.2 Maximum Likelihood

The maximum likelihood (ML) classifier is one of the most popular methods of classification [42]. The goal is to assign the most likely class w_j, from a set of N classes $w_1, \ldots, w_N$, to each feature vector. The most likely class w_j from a given feature vector $\mathbf{x}$ is the one with maximum posterior probability of belonging to the class $P(w_j \mid \mathbf{x})$. Using the Bayes' theorem, we have

$$P(w_j \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid w_j)P(w_j)}{P(\mathbf{x})}$$

On the left side of the equation, there is the *a posteriori* probability of a feature vector $\mathbf{x}$ to belong to the class w_j. On the right side, the *a priori* probability $P(\mathbf{x} \mid w_j)$ that expresses the probability of the feature vector $\mathbf{x}$ being generated by the probability density function of w_j. $P(\mathbf{x})$ and $P(w_j)$ are the *a priori* probability of appearance of feature vector $\mathbf{x}$ and the probability of appearance of each class w_j, respectively.

This model relies on the knowledge of the probability density function underlying each of the classes, as well as the probability of occurrence of the data and the classes. In order to reduce the complexity of such estimations, some assumptions are made. The first assumption generally made is the equiprobability of appearance for each of the feature vector as well as for each of the classes. This assumption reduces the Bayes' theorem to estimate the probability density function for each class:

$$P(w_j \mid \mathbf{x}) = P(\mathbf{x} \mid w_j)$$

Multiple methods can be used to estimate the *a priori* probability. Two of the most widespread methods are the assumption of a certain behavior and the mixture models.

A very common hypothesis is to identify the underlying probability density function with a multivariate normal distribution. In that case the likelihood value is

$$P(\mathbf{x} \mid w_j) = \frac{1}{\sqrt{\Sigma_j}(2\pi)^{n/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu_j)\Sigma_j^{-1}(\mathbf{x}-\mu_j)^T}$$

where Σ_j and μ_j are the covariance matrix and the mean value for class j, respectively. In the case where the determinants of the covariance matrix for each of the classes are equal to each other, the likelihood value becomes the

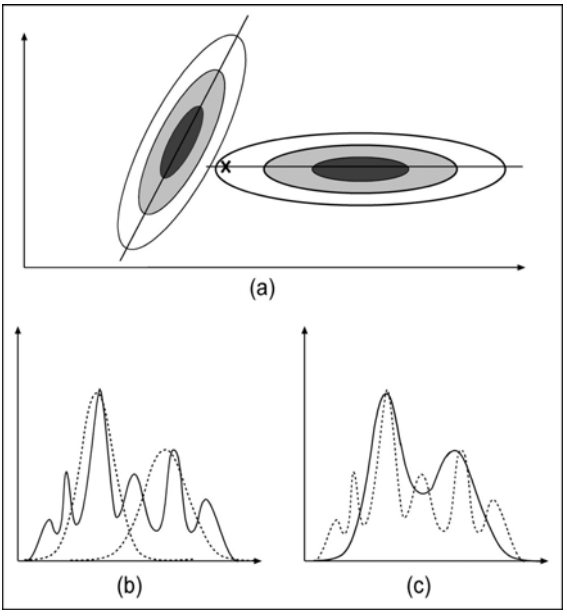


Figure 2.14: (a) Graphic example of the maximum likelihood classification assuming an underlying density model. (b) Unknown probability density function estimation by means of a 2 Gaussian mixture model. (c) Resulting approximation of the unknown density function.

same as the Mahalanobis distances. Figure 2.14(a) shows an example of the effect of this kind of classifier on a sample "X." Although the sample seems to be nearer the left-hand distribution in terms of Euclidean distance, it is assigned to the class on the right hand since the probability of generating the sample is higher than its counterpart.

The other approach is to estimate the model of the probability density function. In the mixture model approach, we assume that the probability density function can be modelled using an ensemble of simple known distributions. If the base distribution is the Gaussian function it is called Gaussian mixture model. The interest in this method consists of the estimation of complex density function using low-level statistics.

The mixture model is composed of a sum of fundamental distributions, following the next expression:

$$p_i(x \mid \Theta) = \sum_{k=1}^{C} p_k(x \mid \theta_k) P_k \tag{2.2}$$

where C is the number of mixture components, P_k is the *a priori* probability of the component k, and θ_k represents the unknown mixture parameters. In our case, we have chosen Gaussian mixture models $\theta_k = \{P_k, \mu_k, \sigma_k\}$ for each set of texture data we want to model. Figures 2.14(b) and 2.14(c) show an approximation of a probability density function with a mixture of two Gaussian and the resulting approximation. Figure 2.14(b) shows the function to be estimated as a continuous line and the Gaussian functions used for the approximation as a dotted line. Figure 2.14(c) shows the resulting approximated function as a continuous line and the function to be estimated as a dotted line as a reference. One can observe that with a determined mixture of Gaussian distributions, an unknown probability density function can be well approximated. The main problem of this kind of approaches resides in its computational cost and the unknown number of base functions needed, as well as the value of their governing parameters. In order to estimate the parameters of each base distribution, general maximization methods are used, such as expectation-maximization (EM) algorithm [42].

However, this kind of techniques are not very suitable as the number of dimensions is large and the training data samples size is small. Therefore, a process of dimensionality reduction is needed to achieve a set of meaningful data. Principal component analysis and Fisher linear discriminant analysis are the most popular dimensionality reduction techniques used in the literature.

2.3.3 Feature Data Dimensionality Reduction

2.3.3.1 Principal Component Analysis

This method is also known as *Karhunen–Loeve* method [38]. Component analysis seeks directions or axes in the feature space that provide an improved, lower dimensional representation of the full data space. The method chooses a dimensionality reducing linear projection that maximizes the scatter of all projected samples. Let us consider a set of M samples $\{x_1, x_2, \ldots, x_M\}$ in an n-dimensional space. We also consider a linear transformation that maps the original space in a lower dimensional space (of dimension m, $m < n$). The new feature vectors y are defined in the following way:

$$y_k = W^T x_k, \quad k = 1, \ldots, M$$

where W is a matrix with orthonormal columns. The total scatter matrix S_T is defined as

$$S_T = \sum_{k=1}^{M} (x_k - \mu)(x_k - \mu)^T$$

where M is the number of samples and μ is the mean vector of all samples. Applying the linear transformation W^T, the scatter of the transformed feature vectors is $W^T S_T W$. PCA is defined as to maximize the determinant of the scatter of the transformed feature vectors:

$$W_{\text{opt}} = \text{argmax}|W^T S_T W| = [w_1 w_2 \cdots w_m]$$

where $\{w_i \mid i = 1, 2, \ldots, m\}$ is the set of n-dimensional eigenvectors of S_T corresponding to the m largest eigenvalues.

Therefore, PCA seeks the directions of maximum scatter of the input data, which correspond to the eigenvectors of the covariance matrix having the largest eigenvalues. The n-dimensional mean vector μ and the $n \times n$ covariance matrix Σ are computed for the full dataset.

In summary, the eigenvectors and eigenvalues are computed and sorted in decreasing order. The k eigenvectors having the largest eigenvalues are chosen. With those vectors a $n \times m$ matrix W_{opt} is built. This transformation matrix defines an m-dimensional subspace. Therefore, the representation of the data onto this m-dimensional space is

$$y = A^t(x - \mu)$$

PCA is a general method to find the directions of maximum scatter of the set of samples. This fact however does not ensure that such directions will be optimal for classification. In fact, it is well known that some specific distributions of the samples of the classes result in projection directions that deteriorate the discriminability of the data. This effect is shown in Fig. 2.15 in which the loss of information when projecting to the PCA direction clearly hinders the discrimination process. Note that both projections of the clusters on the PCA subspace overlap.

2.3.3.2 Fisher Linear Discriminant Analysis

A classical approach to find a linear transformation that discriminates the clusters in an optimal way is discriminant analysis. Fisher linear discriminant

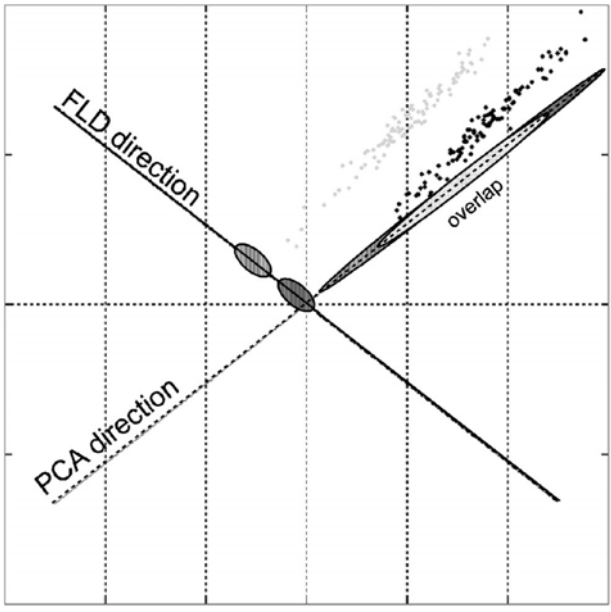


Figure 2.15: Example of the resulting direction using principal component analysis (PCA) and Fisher linear discriminant (FLD).

analysis [38, 39] seeks a transformation matrix W such that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter S_{B} be defined as follows:

$$S_{\mathrm{B}} = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{2.3}$$

where μ_i is the mean value of class X_i, μ is the mean value of the whole data, c is the number of classes, and N_i is the number of samples in class X_i. Let the within-class scatter be

$$S_{\mathrm{W}} = \sum_{i=1}^{c} \sum_{x_{k,i} \in X_i} (x_{k,i} - \mu_i)(x_{k,i} - \mu_i)^T \tag{2.4}$$

where μ_i is the mean value of class X_i, c is the number of classes, and N_i is the number of samples in class X_i. If S_{W} is not singular, the optimal projection matrix W_{opt} is chosen as the matrix that maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant

of the within-class scatter matrix of the projected samples:

$$W_{\text{opt}} = \text{argmax}_W \frac{\left|W^T S_{\text{B}} W\right|}{\left|W^T S_{\text{W}} W\right|} = [\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w}_m] \tag{2.5}$$

where $\mathbf{w}_i$, $i = 1 \ldots m$, is the set of S_{W}-generalized eigenvectors of S_{B} corresponding to the m largest generalized eigenvalues.

Opposed to PCA behavior, Fisher linear discriminant (FLD) emphasizes the direction in which both classes can be better discriminated. FLD uses more information about the problem as the number of classes and the samples in each of the classes must be known *a priori*. In Fig. 2.15 the projections on the FLD subspace are well separated.

In real problems, it can occur that it is not possible to find an optimal classifier. A solution is presented by assembling different classifiers.

2.3.4 AdaBoost Procedure

Adaptative Boosting (AdaBoost) is an arcing method that allows the designer to continue adding "weak" classifiers until some desired low-training error has been achieved [40,41]. A weight is assigned to each of the feature points, these weights measure how accurate the feature point is being classified. If it is accurately classified, then its probability of being used in subsequent learners is reduced or emphasized otherwise. This way, AdaBoost focuses on difficult training points.

Figure 2.16 shows a diagram of the general process of boosting. The input data is resampled according to the weights of each feature data. The higher the weight the most probable the feature point will be in the next classification. The new set of feature points are inputs of the new classifier to be added to the process. At the end of the process, the responses of all the classifiers are combined to form the "strong" classifier.

AdaBoost is capable of performing a feature selection process while training. In order to perform both tasks, feature selection and classification process, a weak learning algorithm is designed to select the single features that best separate the different classes. For each feature, the weak learner determines the optimal classification function, so that the minimum number of feature points is misclassified. The algorithm is described as follows:

- Determine a supervised set of feature points $\{x_i, c_i\}$ where $c_i = \{-1, 1\}$ is the class associated to each of the feature classes.

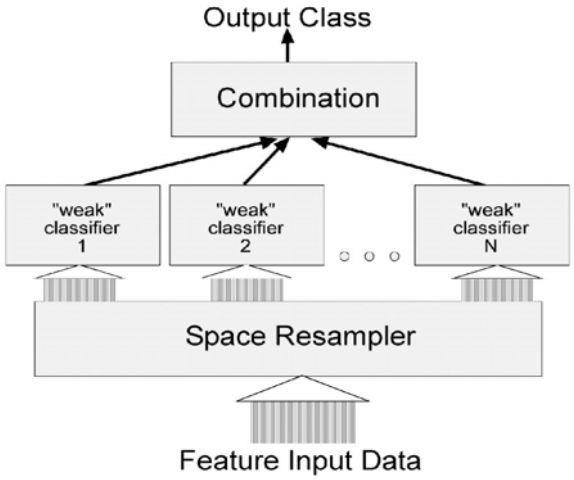


Figure 2.16: Block diagram of the AdaBoost procedure.

- Initialize weights $w_{1,i} = \frac{1}{2}m, \frac{1}{2l}$ for $c_i = \{-1, 1\}$ respectively, where m and l are the number of feature points for each class.
- For $t = 1..T$
  - Normalize weights

    $$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,i}}$$

    so that w_t is a probability distribution.
  - For each feature, j train a classifier, h_j which is restricted to using a single feature. The error is evaluated with respect to w_t, $\epsilon_j = \sum_i w_i |h_j(x_i) - c_i|$.
  - Choose the classifier, h_t with the lowest error ϵ_t.
  - Update the weights:

    $$w_{t+1,i} = w_{t,i}\beta_t^{e_i}$$

    where $e_i = 1$ for each well-classified feature and $e_i = 0$ otherwise. $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$. Calculate parameter $\alpha_t = -\log(\beta_t)$.
- The final strong classifier is

$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

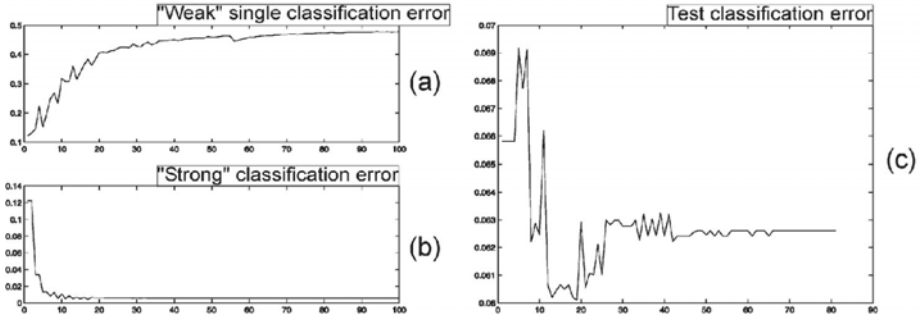


Figure 2.17: Error rates associated to the AdaBoost process. (a) Weak single classification error. (b) Strong classification error on the training data. (c) Test error rate.

Therefore, the strong classifier is the ensemble of a series of simple classifiers, $h_t(x)$, called "weaks". Parameter α_t is the weighting factor of each of the classifiers. The loop ends when the classification error of a weak classifier is over 0.5, the estimated error for the whole strong classifier is lower than a given error rate or if we achieve the desired number of weaks. The final classification is the result of the weighted classifications of the weaks. The process is designed so that if $h(x) > 0$, then pixel x belongs to one of the classes.

Figure 2.17 shows the evolution of the error rates for the training and the test feature points. Figure 2.17(a) shows the error evolution of each of the weak classifiers. The abscise axis is the number of the weak classifier, and the ordinate axis is the error percentage of a single weak. The figure illustrates how the error increases as more weak classifiers are added. This is because each new weak classifier focusses on the misclassified data of the overall system. Figure 2.17(b) shows the error rate of the system response on the training data. The abscise axis represents the number of iterations, that is, the number of classifiers added to the ensemble. As it is expected, the error rate decreases to very low values. This, however, does not ensure a test classification error of such accuracy. Figure 2.17(c) shows the test error rate. One can observe, that the overall error has a decreasing tendency as more weak classifiers are added to the process.

Therefore, the weak classifier has a very important role in the procedure. Different approaches can be used; however, it is relatively interesting to center our attention in low-time-consuming classifiers.

The first and the most straight forward approach to a weak is the perceptron. The perceptron is constituted by a weighed sum of the inputs and an adaptative

threshold function. This scheme is easy to embed in the adaboost process since it relies on the weights to make the classification.

Another approach to be taken in consideration is to model the feature points as Gaussian distributions. This allows us to define a simple scheme by simply calculating the weighed mean and weighed covariance of the classes at each step t of the process:

$$\mu_{i,t}^{j} = \sum_{i} w_{i,t} x_i \qquad \Sigma_{i,t}^{j} = \sum_{i} w_{i,t} \left(x_i - \mu_{i,t}^{j}\right)^2$$

for each x_i^j point in class C_j. $W_{i,j}$ are the weights for each data point.

If feature selection is desired, this scheme is highly constrained to the N features of the N-dimensional feature space. If N is not enough large, the procedure could not improve its performance.

Both, the feature extraction and the classification processes, are the central parts of the tissue characterization framework. Next section is devoted to explain the different frameworks where these processes are applied for tissue characterization of IVUS images as well as provide quantitative results of their performance.

2.4 Results and Conclusions

The goal of automatic tissue characterization is to identify the different kind of plaques in IVUS images. This process requires two tasks: identification of what the plaque is and labelling of the different areas of the plaque. Figure 2.18 illustrates roughly the procedure of supervised tissue characterization. The IVUS image (see Fig. 2.18(a)) is preprocessed and sent to the automatic tissue characterization system. Figure 2.18(b) illustrates the physicians-assisted segmentation of IVUS, which will constitute a part of the training dataset. Figure 2.18(c) shows the first step, the accurate location of the lumen–plaque border and the adventitia border. Between both borders, the plaque is the region of interest to be classified. Figure 2.18(d) exemplifies the tissue characterization process. We are focussed on the plaque, and try to find and label areas corresponding to different plaques. In the figure, light gray areas are soft plaque, white areas are fibrotic plaque, and dark gray areas are calcium plaque. At the end of the process we obtain the tissue characterized IVUS (Fig. 2.18(e)) to be used by

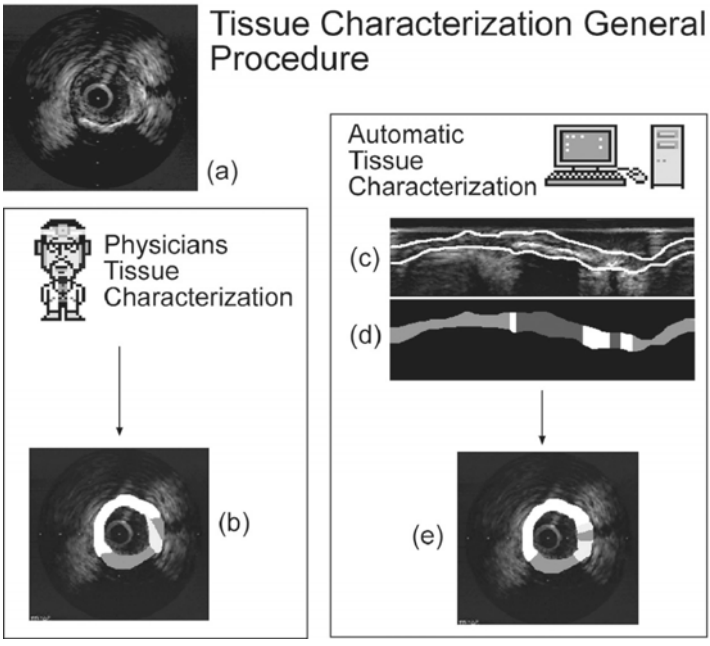


Figure 2.18: The tissue characterization process can be done manually by physicians (a) and (b) or by an automatic process (a), (c), (d), and (e). (c) Segmentation of the lumen–plaque border and adventitia border. (d) Processed image of the plaque characterization. (e) Final IVUS tissue characterization.

the physicians. These results have been used to validate against the manually segmented plaque regions (Fig. 2.18(b)).

Therefore, though we are concerned with tissue characterization, we cannot forget the segmentation of the plaque. A brief review on how to segment the plaque is exposed in the next section.

2.4.1 Segmentation of the Plaque

The segmentation of the plaque is a really important step before tissue characterization. There are multiple ways to achieve this goal [19, 20, 43–45]. In particular, we will focus on two general lines of work: the line proposed in [43] and the line proposed in [19,20].

In [43] the process of segmentation relies on a manual definition of a region of interest. Using that region of interest a Sobel-like edge operator with

neighborhoods of 5×5 and 7×7 is applied. Once we have these features extracted, the problem of identification of the vessel wall and plaque border is solved by finding the optimal path in a two-dimensional graph. The key of the graph searching is to find the appropriate cost functions. In the paper, the authors propose different cost functions depending on whether the lumen–plaque border or the adventitia border is desired.

Having in mind the tissue classification goal of the process, in [19, 20], the authors try to find a segmentation of the overall tissue independent of what kind of tissue it is, to distinguish the lumen–plaque border. Therefore, the method consists of selecting a feature space and a classifier. This method takes advantage of the fact that for tissue characterization the same scheme must be used. Thus, a classifier is trained for general tissue discrimination. Hence, in the overall process the feature extraction process is performed once for both, plaque segmentation and tissue identification. What is different in both approaches is the classification selection and training data, and the post processing steps.

The classification step is performed using a fast classifier, boosting methods, or ML. The result of this step is a series of unconnected areas that are related to tissues. In order to find the exact location of the lumen–plaque border, a fast parametric snake is let to deform over the unconnected areas [46]. The snake performs a double task: first, it finds a continuous boundary between blood and plaque. The second task is that it ensures an interpolation and a fill-in process in regions where tissue is not located or not reliable (such as areas with reverberations due to the guide-wire, etc). The adventitia border is found by context using a 5×5 Sobel-like operator and deformable models. Figure 2.19 shows an example of a possible scheme for border location. First the IVUS image is transformed to cartesian coordinates. Then, a texture feature extraction step is performed. A classification scheme is trained to distinguish blood from tissue. At the end, a snake deforms to adapt to the classified IVUS image and locates accurately the blood–plaque border.

2.4.2 Tissue Characterization

2.4.2.1 Methodology

We begin our process of tissue characterization taking the IVUS image and transforming it to cartesian coordinates (Fig. 2.20(a)). Once the cartesian

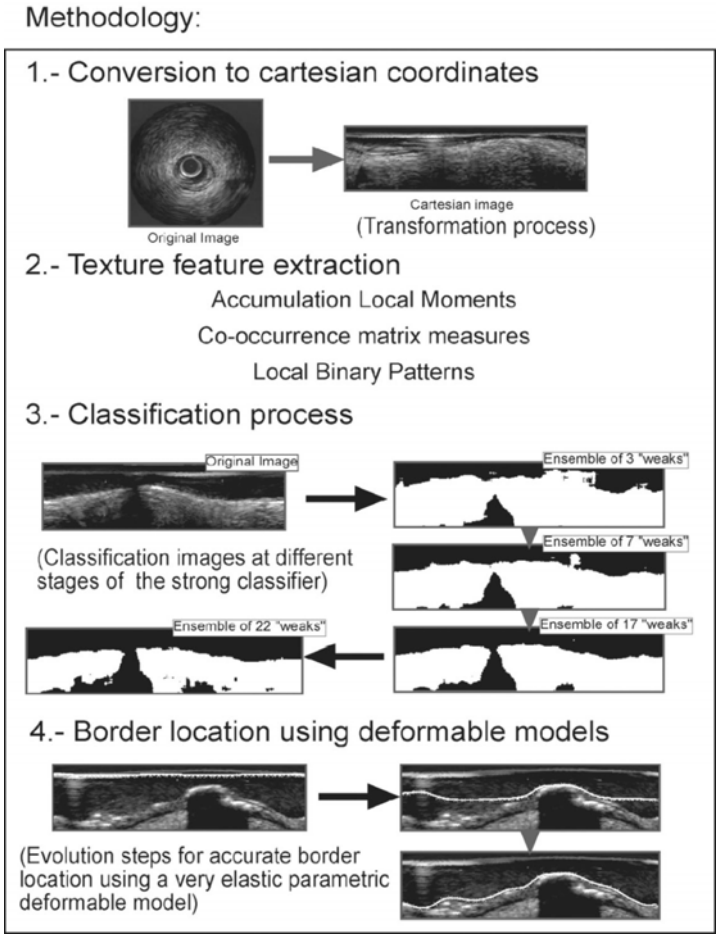


Figure 2.19: Plaque segmentation (see text).

transformation is done, artifacts are removed from the image (Fig. 2.20(b)). There are three main artifacts in an IVUS image: the transducer wrapping, which creates a first halo at the center of the image (in the cartesian image the echo is shown at the top of the image); the guide-wire effect, which produces an echo reverberation near the transducer wrapping; and the calibration grid, which are markers at a fixed location that allow the physicians to evaluate quantitatively the morphology and the lesions in the vessel. With the artifacts removed, we proceed to identify intima and adventitia using the process described in the former section. At this point, we have the plaque located and we are concerned with tissue identification (Fig. 2.20(c)). The tissue classification process is divided

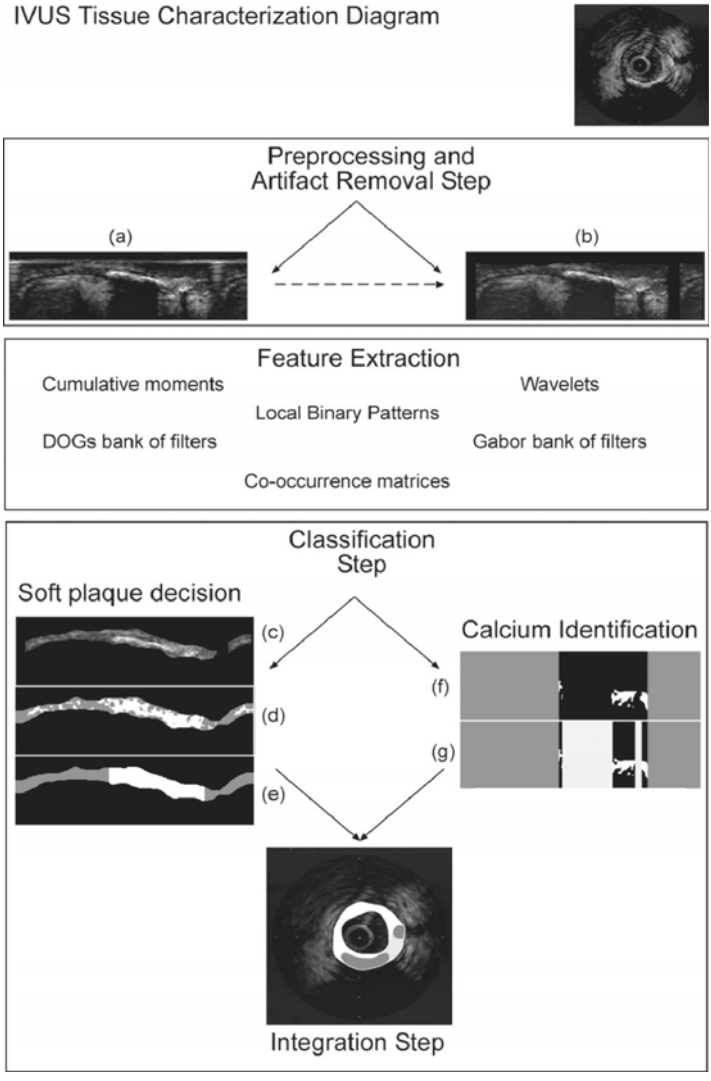


Figure 2.20: Tissue characterization diagram.

in three stages: First, the soft–hard classification (Figs. 2.20(d) and 2.20(e)), in which the soft plaque, the hard plaque, and calcium are separated. In the second stage, the calcium is separated from the hard plaque (Figs. 2.20(f) and 2.20(g)). At the last stage, the information is fused and the characterization is completed. We will refer later to this diagram to explain some parts of the process. Recall that the plaque is the area comprised between the intima and the adventitia. With both borders located we can focus on the tissue of that area.

For such task, the three stages scheme formerly described is used. Regarding the first stage of the process, a classification is performed on the feature space. At this point, a feature space and a classifier must be selected. To help to choose which feature space and which classifier to use, we try each of the feature spaces with a general purpose classifier, the k-nearest neighbors method used as a ground truth classifier. Regardless the classifier used, the information provided at the output of the system is a pixel classification. Using these data we can further process the classification result incorporating region information from the classification process and obtain clear and smoother borders of the soft and the mixed plaques. Different processes can be applied to achieve this goal, two possible approaches are region-based area filtering and classification by density filtering. In a region-based area filtering the less significant regions in terms of size are removed from the classification. On the other hand, the other method relies on keeping the regions that have high density of classification responses. As the classification exclusively aims to distinguish between soft and hard plaque, a separate process is added to separate hard plaque from calcium.

Once soft and hard plaque are distinguished, we proceed to identify what part of the hard plaque corresponds to calcium. One can argue why not to include a third class in the previous classifier. The reason we prefer not to do so is because experts' identification of calcium plaques is performed by context. Experts use the shadowing produced by the absorbtion of the echoes, behind a high-echoreflective area, to label a certain area as calcium. In the same way, we take the same approach. On the other hand, the fact of including a third class only hinders the decision process and increases the classifier complexity. Therefore, the calcium identification process is made by finding the shadowing areas behind hard plaque. Those areas are easily identified because the soft–hard classification also provides this information (Fig. 2.20) since shadowing areas are classified as nontissue. We can see a plausible way of finding calcified areas. Figure 2.20(f) shows the classification result under the adventitia border of the "hard" tissue. Dark gray level areas are regions with soft plaque and, therefore, do not provide information of the calcium composition of the plaque. We use one of the previous classified images, the soft–hard classification or the blood–plaque classification. In white, it is displayed the regions of tissue under the adventitia border in the area of interest. Figure 2.20(g) shows in light gray the areas of shadowing, and therefore, the areas labelled as calcium.

To end the process, the last stage is devoted to recast the resulting classification to its original polar domain by means of a simple coordinate transformation.

2.4.2.2 Experimental Results

To evaluate the results, a classification of over 200 full-tissue regions from 20 different patients has been performed. The training set is a subset of two thirds of the overall data determined using the bootstrapping strategy. The rest of the data has been used as test set. Previously, different physicians have determined and delineated plaque regions in each full-tissue image.

The first experiment is to set a ground truth for the feature spaces, as a measure to evaluate their description power. We have used k-nearest neighbors as a ground truth classifier. To choose the number of neighbors, we select a feature space and evaluate the performance for different values of k. Tables 2.2, 2.3 and 2.5 show the error rates for pixel classification (RAW Error) and postprocessed classification taking into account neighboring information and density of classifier cluster responses (Post Error). These tables also show the percentages of false positives (FP) and false negatives (FN) for both errors. The FP and FN are included as they give information of the possible geometry of samples in the feature space.

Table 2.2 illustrates the results regarding the selection of the number of neighbors k. It can be seen that for $k = 7$ a lower pixel error rate is obtained. Therefore, the performance of the feature spaces will be evaluated using 7-nearest neighbors. The result of the classification of the test data using all feature spaces and 7-nearest neighbors classifier is shown in Table 2.3. Observing the RAW data error rate, the best overall feature spaces are the co-occurrence matrices, local binary patterns, derivatives of Gaussian, and accumulation local moments. These results are confirmed looking at the postprocessing error rate and ratifies

Table 2.2: Selection of the parameter k, using local binary pattern feature space as a reference

k value	RAW error	FP	FN	Post error	FP	FN
3	33.94	25.13	8.80	10.16	3.46	6.69
7	25.67	9.67	16.23	3.45	2.67	0.81
15	32.93	26.19	6.74	5.81	3.46	2.34

Table 2.3: Feature space performance discriminating hard plaque from soft plaque using k-nearest neighbor

Feature space	RAW error	FP	FN	Post error	FP	FN
Co-occurrence measures	22.36	10.91	11.45	10.88	4.19	6.68
Derivative of Gaussian	27.81	23.51	4.95	16.29	16.67	0.04
Gabor filters	35.26	18.86	17.22	16.26	16.49	0.07
Wavelets	45.05	20.52	24.90	31.78	24.40	7.68
Accumulation local moments	31.72	16.42	15.30	12.17	11.36	0.81
Local binary patterns	25.67	9.67	16.23	3.45	2.67	0.81

the qualitative evaluation shown in Table 2.4, where we observe that the same feature spaces are the ones that perform best. Analyzing each of the feature spaces in terms of FP and FN rates, we can deduce that Co-occurrence feature space has good discrimination power, having a "symmetric" nature where both FP and FN rates are comparable. In the same sense, we can deduce that the overlapping of both classes is similar. Derivatives of Gaussian's filter space have tendency to over-classify hard plaque. The classes in the feature space are not very well defined as hard plaque must have a higher scatter than the soft plaque. Gabor filter's bank gives a good description of both classes as they have similar false rates. However, both classes are very overlapped giving a hard time to the classification process. Wavelets overlapping of classes in the feature space is extremely high; therefore, it describes bad each of the classes. Accumulation local moments have similar description power than Gabor filter's bank; however, the different responses from both allow a much better postprocessing in accumulation local moments. This fact allows us to suppose that the classification error points in the image domain are much more scattered and

Table 2.4: Descriptive table of the discriminative power of each feature space using k-nearest neighbors

Feature space	Qualitative speed	Qualitative performance
Co-occurrence measures	Slow	Good
Gabor space	Slow	Acceptable
Wavelets	Fast	Poor
Derivative of Gaussian	Slow	Acceptable
Accumulation local moments	Fast	Good
Local binary patterns	Fast	Good

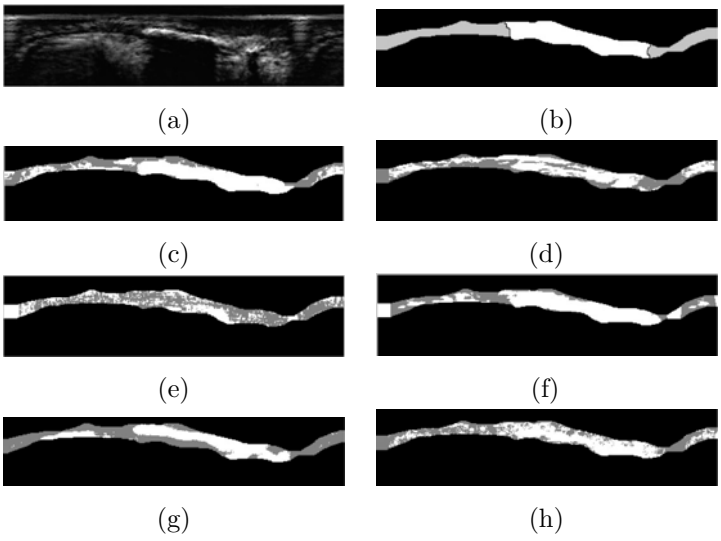


Figure 2.21: Tissue pixel classification data using 7-nearest neighbors method on different feature spaces. (a) Original image in cartesian coordinates. (b) Expert manual classification of tissue. (c) Co-occurrence feature space. (d) Gabor feature space. (e) Wavelets feature space. (f) Derivative of Gaussian feature space. (g) Accumulation local moments feature space. (h) Local binary patterns feature space.

with very few local density. Local binary patterns have good descriptive power as well as giving a more sparse pattern in FP and FN in the image domain. Figure 2.21 provides a graphical example of the performance of 7-nearest neighbors method applied to several feature spaces. Observing the images, we realize that scale-space processes, derivative of Gaussian, Gabor filters, and wavelets have poor to acceptable discrimination power, and therefore, are not suitable for the task of tissue discrimination. On the other hand, statistic-based feature spaces and structure feature spaces have acceptable to good performances. Table 2.4 details the conclusions arisen from the experiment. The qualitative speed nomenclature (fast/slow) indicates the viability of the feature space technique to be included in a real time or near-real time process. A “fast” scheme denotes a method over 10 times faster than the “slow” one. Because the results are obtained using prototypes and not a full application, no absolute time measure is provided. Note, also, that the images displayed are pixel-based classification results and have no further smoothing postprocessing. To further develop our

Table 2.5: Feature space performance using FLD and Mahalanobis distance

Feature space	RAW error	FP	FN	Post error	FP	FN
Co-occurrence measures	40.88	34.66	6.20	12.91	12.10	0.81
Accumulation local moments	35.50	20.34	16.16	13.83	13.02	0.81
Local binary patterns	26.37	5.76	20.85	6.93	1.52	5.47

discussion we will only take the three best postprocessed data performing feature spaces: co-occurrence matrix measures, accumulation local moments, and local binary patterns. Up to this point we have neither taken into account complexity of the methods nor time issues. However, these are critical parameters in real applications, thus, we consider them in our following discussions.

Once the feature space is selected, the next decision is to find the most suitable classifier taking into account our problem constraints, if any. We are concerned with speed issues, therefore, simple but powerful classifiers are required. Because the high dimensionality of two of the feature spaces selected (co-occurrence matrix measures have about 24 features per distance and accumulation local moments have 81 features) a dimensionality reduction step is desired. PCA is the first obvious choice, but because great amount of overlapping data the results are worse than using Fisher's linear discriminant analysis which is focalized in finding the most discriminative axes for our given set of data. The result of this experiment is shown in Table 2.5. We use maximum likelihood combined with a Fisher linear discriminant analysis reduction. As local binary patterns do not need dimensionality reduction due to the small amount of features computed (three features), the comparison with this method is done by just classifying with the ML method. As expected, the raw results are much worse with this kind of classifier. Co-occurrence matrix measures take the worse part doubling their error rate. However, local binary patterns, though they have also worse error rate with ML, manage to be the most discriminative of the three methods. This fact is also shown in the postprocessing, where local binary patterns still have the lower error ratio. Co-occurrence matrix measures regain their discrimination power after the postprocessing.

Therefore, using one of the fastest classifiers, ML, one achieves, at least, a classification ratio over 87% (with accumulation local moments). If the selected feature space is local binary patterns, the scheme is the fastest possible scheme as local binary patterns are computationally efficient and low-time consuming as

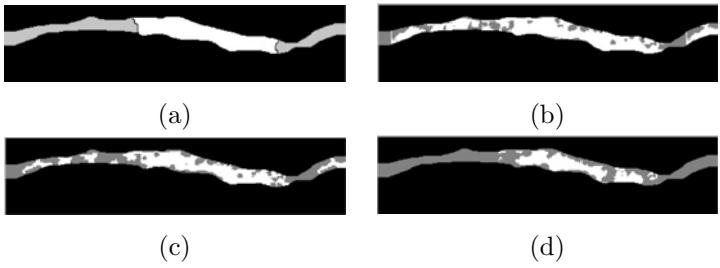


Figure 2.22: Boosting procedure for tissue characterization at different stages of its progress. (a) Expected hand classification by an expert. (b) First stage of the boosting procedure. (c) Classification with a five classifiers ensemble. (d) Classification with 10 "weaks" ensemble.

well as the ML classifier does not transform data in another feature space. This scheme is really well suited for real-time or near-real-time applications because of both time efficiency and reliability in the classification. This is, however, by no means the only near-real-time configuration available since accumulation local moments are computationally as fast as local binary patterns. However, the FLD dimensionality reduction hinders the process due to the complexity of the data in its original feature space. To overcome this problem, other classifiers can be used. The necessity to find reliable and fast classifiers lead us to boosting techniques. Boosting techniques allow a fast and simple classification procedure to improve its performance as well as maintaining part of its speed. To illustrate this fact Fig. 2.22 shows the evolution of the classification when more classifiers are added to the *strong classifier*. Figure 2.22(a) shows the expected hand classification by a physician. Figure 2.22(b) shows the base classification of a single "weak". Figure 2.22(c) illustrates the result of the classification using an ensemble of five classifiers. Figure 2.22(d) shows the resulting classification after the addition of 10 weak classifiers to the ensemble. The error rates at different stages of the process are also shown in Table 2.6. These results are computed using a ML method as a *weak classifier* on the accumulation local moments space. The numbers show how the error rate is improved, and, though the raw classification error rate is nearly immutable, we can observe that there is a great change in the classification data points distribution in the image domain since the FP and FN rates drastically change. The postprocessing error rate gives better description of what is happening. The disposition of the error points in the classification image is more sparse and unrelated to their neighborhood, allowing better

Table 2.6: Error rates using boosting methods with maximum likelihood with the accumulation local moments space

Ensemble no.	RAW error	FP	FN	Post error	FP	FN
Base error	34.86	28.20	6.98	41.94	40.33	1.10
Ensemble of 5 c.	29.38	16.32	13.32	33.17	31.87	1.10
Ensemble of 10 c.	31.44	7.36	23.37	7.92	3.22	4.76

postprocessing and classification rates. In this case, the classification rate is over 92% with a classifier as fast as applying 10 times a threshold. Therefore, using accumulation local moments and boosting techniques we have another fast and highly accurate scheme for real-time or near-real-time tissue characterization.

Up to this point, we have discussed the reliability of the soft plaque versus hard plaque discrimination process, which is our main concern, since the identification of calcium is reduced to the part of hard plaque with a large shadowing area. Using the method described in the former section, 99% of the calcium plaque is correctly identified. Figure 2.23 shows some results of the tissue characterization process. Figures 2.23(a) and 2.23(b) show the characterization of a soft plaque. In Figs. 2.23(c) and 2.23(d), there are two different kind of plaques detected, calcium (gray region) and soft plaque (white region). Figures 2.23(e) and 2.23(f) show the characterization of the three kind of plaques: fibrotic (light gray region), soft plaque (white region), and calcium (dark gray region).

2.4.3 Conclusions

Tissue characterization in IVUS images is a crucial problem for the physicians for studying the vascular diseases. However, this task is complex and suffers from multiple drawbacks (slow manual process, subjective interpretation, etc.) Therefore, automatic plaque characterization is a highly desirable tool.

However, automatic tissue characterization is a problem of high complexity. First of all, we need a unique and powerful description of the tissues to be classified. This is done by the feature extraction process, that in order to obtain complete and meaningful description, image features should be based on texture. Thus, a study of the most representable feature spaces is done, to conclude with some enlightening results. After analyzing the experimental results, we conclude that co-occurrence matrix measures, local binary patterns, and

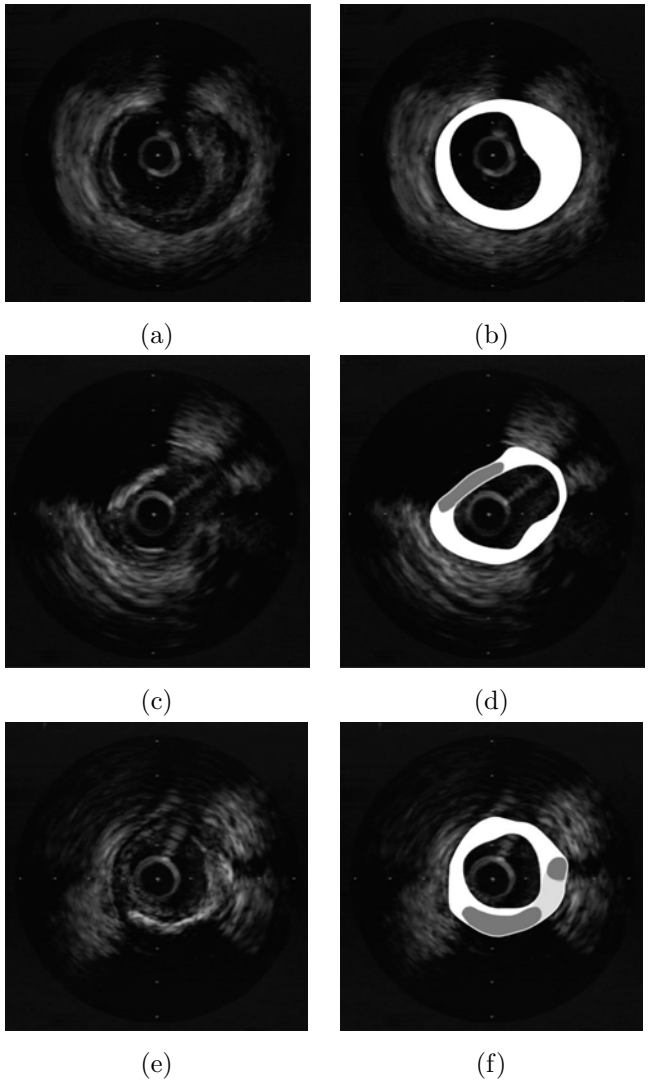


Figure 2.23: Tissue characterization results: (b), (d), and (f) White labels soft plaque, dark gray areas are displayed where calcium plaques are located, and light gray areas labels hard plaque. (a), (c), and (e) Original images.

accumulation local moments are good descriptors of the different kind of plaque tissues. However, local binary patterns and accumulation local moments are also fast, in terms of low-time processing. On the other hand, the classification of the feature data is a critical step. Different approaches to the classification problem are described and proposed as candidates in our framework. We proved that

k-nearest neighbor method gives the best performance as a classifier. But, ML and methods based on an ensemble of classifiers have high discrimination rate and lower classification time. Therefore, two real-time or near-real-time approaches are proposed: The first method combines local binary patterns with ML methods. The second method uses accumulation local moments and boosting techniques.

In conclusion, we have presented a general fully automatic and real-time or near-real-time framework with high accuracy plaque recognition rate for tissue characterization in IVUS images.

Questions

1. *What is tissue characterization in IVUS images?*
2. *Why is automatic tissue characterization an important issue?*
3. *Why do we use texture-based descriptors?*
4. *Why do we use supervised classification?*
5. *What is the feature space?*
6. *Why is dimensionality reduction needed in the classification process?*
7. *What is the main idea under the boosting classification?*
8. *What is the segmentation of the plaque for?*
9. *Discuss the methodology for the tissue characterization framework.*
10. *Which are the most reliable frameworks for real-time classification?*

Bibliography

[1] Wickline, S., Beyond intravascular imaging: Quantitative ultrasonic tissue characterization of vascular pathology, In: IEEE Ultrasonics symposium, 1994, pp. 1589–1597.

[2] Arul, P. and Amin, V., Characterization of beef muscle tissue using texture analysis of ultrasonic images, In: Proceedings of the Twelfth Southern Biomedical Engineering Conference, 1993, pp. 141–143.

[3] Mojsilovic, A. and Popovic, M., Analysis and characterization of myocardial tissue with the wavelet image extension [US images], In: Image Processing, 1995 (Proceedings) Vol. 2, pp. 23–26, 1995.

[4] Jin, X. and Ong, S., Fractal characterization of kidney tissue sections, In: Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers, Proceedings of the 16th Annual International Conference of the IEEE, Vol. 2, pp. 1136–1137, 1994.

[5] Cohen, F. and Zhu, Q., Quantitative soft-tissue characterization in human organs using texture/attenuation models, In: Proceedings in Multidimensional Signal Processing Workshop, 1989, pp. 47–48.

[6] Mavromatis, S. and Boi, J., Medical image segmentation using texture directional features, In: Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE, Vol. 3, pp. 2673–2676, 2001.

[7] Mavromatis, S., Mammographic mass classification using textural features and descriptive diagnostic data, In: Digital Signal Processing, DSP 2002. 14th International Conference on, Vol. 1, pp. 461–464, 2002.

[8] Donohue, K. and Forsberg, F., Analysis and classification of tissue with scatterer structure templates, IEEE Trans. Ultrasonics, Ferroelect. Frequency Control, Vol. 46, No. 2, pp. 300–310, 1999.

[9] Ravizza, P., Myocardial tissue characterization by means of nuclear magnetic resonance imaging, In: Computers in Cardiology 1991 (Proceedings), pp. 501–504.

[10] Vandenberg, J., Arterial imaging techniques and tissue characterization using fuzzy logic, In: Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, 1994, pp. 239–243.

[11] Nailon, W. and McLaughlin, S., Comparative study of textural analysis techniques to characterize tissue from intravascular ultrasound, In: Proceedings of the IEEE International Conference of Image Processing, Switzerland, IEEE Signal Processing Society, Piscataway, NJ, 1996, pp. 303–305.

[12] Nailon, W. and McLaughlin, S., Intravascular ultrasound image interpretation, In: Proceedings of the International Conference on Pattern Recognition, Austria, IEEE Computer Society Press, Los Alamitos, CA, 1996, pp. 503–506.

[13] Nailon, W., Fractal texture analysis: An aid to tissue characterization with intravascular ultrasound, In: Proceedings 19th International Conference, IEEE/EMBS, 1997, pp. 534–537.

[14] Spencer, T., Characterization of atherosclerotic plaque by spectral analysis of 30mhz intravascular ultrasound radio frequency data, In: IEEE Ultrasonics Symposium, 1996, pp. 1073–1076.

[15] Dixon, K., Characterization of coronary plaque in intravascular ultrasound using histological correlation, In: 19th International Conference, IEEE/EMBS, pp. 530–533, 1997.

[16] Ahmed, M. and Leyman, A., Tissue characterization using radial transform and higher order statistics, In: Nordic Signal Processing Symposium, 2000, pp. 13–16.

[17] de Korte, C. L. and van der Steen, A. F. W., Identification of atherosclerotic plaque components with intravascular ultrasound elastography in vivo: A yucatan pig study, Circulation, Vol. 105, No. 14, pp. 1627–1630, 2002.

[18] Zhang, X. and Sonka, M., Tissue characterization in intravascular ultrasound images, IEEE Trans. Med. Imaging, Vol. 17, No. 6, pp. 889–899, 1998.

[19] Pujol, O. and Radeva, P., Automatic segmentation of lumen in intravascular ultrasound images: An evaluation of texture feature extractors, In: Proceedings for IBERAMIA, 2002, pp. 159–168.

[20] Pujol, O. and Radeva, P., Near real time plaque segmentation of ivus, In: Proceedings of Computers in Cardiology, 2003, pp. 159–168.

[21] Randen, T. and Husoy, J. H., Filtering for texture classification: A comparative study, Pattern Recogn., Vol. 21, No. 4, pp. 291–310, 1999.

[22] Haralick, R., Shanmugam, K., and Dinstein, I., Textural features for image classification, IEEE Trans. System, Man, Cybernetics, Vol. 3, pp. 610–621, 1973.

[23] Tuceryan, M., Moment based texture segmentation, Pattern Recogn. Lett., Vol. 15, pp. 659–668, 1994.

[24] Lindeberg, T., Scale-Space Theory in Computer Vision, Kluwer, Dordrecht, Netherlands, 1994.

[25] Jain, A. and Farrokhnia, F., Unsupervised texture segmentation using gabor filters, In: Systems, Man and Cybernetics, 1990 (Conference Proceedings), pp. 14–19.

[26] Mallat, S., A theory for multiresolution signal decomposition: The wavelet representation, IEEE Trans. Pattern Anal. Machine Intell., Vol. 11, No. 7, pp. 674–694, 1989.

[27] Mandelbrot, B., The Fractal Geometry of Nature, W. H. Freeman, New York, 1983.

[28] Ojala, T., Pietikainen, M., and Maenpaa, T., Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Machine Intell., Vol. 24, No. 7, pp. 971–987, 2002.

[29] Julesz, B., Visual pattern discrimination, IRE Trans. Inf. Theory, Vol. IT-8, pp. 84–92, 1962.

[30] Ohanian, P. and Dubes, R., Performance evaluation for four classes of textural features, Pattern Recogn., Vol. 25, No. 8, pp. 819–833, 1992.

[31] Martinez, J. and Thomas, F., Efficient computation of local geometric moments, IEEE Trans. Image Process., Vol. 11, No. 9, pp. 1102–1111, 2002.

[32] Caelli, T. and Oguztoreli, M. N., Some tasks and signal dependent rules for spatial vision, Spatial Vision, No. 2, pp. 295–315, 1987.

[33] Chaudhuri, B. and Sarkar, N., Texture segmentation using fractal dimension, IEEE Trans. Pattern Anal. Machine Intell., Vol. 17, No. 1, pp. 72–77, 1995.

[34] Lindeberg, T., Scale-space theory: A basic tool for analysing structures at different scales, J. Appl. Stat., Vol. 21, No. 2, pp. 225–270, 1994.

[35] Rao, R. and Ballard, D., Natural basis functions and topographic memory for face recognition, In: Proceedings of International Joint Conference on Artificial Intelligence, 1995, pp. 10–17.

[36] Lumbreras, F., Segmentation, Classification and Modelization of Textures by means of Multiresolution Descomposition Techniques, PhD Thesis, Computer Vision Center, Universitat Autònoma de Barcelona, 2001.

[37] Jain, A. and Farrokhnia, F., A multi-channel filtering approach to texture segmentation, In: Proceedings of Computer Vision and Pattern Recognition, CVPR'91, 1991, pp. 364–370.

[38] Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, New York, 1971.

[39] Belhumeur, P., Eigenfaces vs fisherfaces: Recognition using class specific linear projection, IEEE Pattern Analy. Machine Intell., Vol. 19, No. 7, pp. 711–720, 1997.

[40] Schapire, R. E., The boosting approach to machine learning. An overview, In: MSRI Workshop on Nonlinear Estimation and Classification, 2002.

[41] Viola, P. and Jones, M., Rapid object detection using a boosted cascade of simple features, In: Conference on Computer Vision and Pattern Recognition, 2001, pp. 511–518.

[42] Duda, R. and Hart, P., Pattern Classification, Wiley InterScience, New York, 2001. Second Edition.

[43] Sonka, M. and Zhang, X., Segmentation of intravascular ultrasound images: A knowledge-based approach, IEEE Trans. Med. Imaging, Vol. 17, No. 6, pp. 889–899, 1998.

[44] von Birgelen, C., Computerized assessment of coronary lumen and atherosclerotic plaque dimensions in three-dimensional intravascular ultrasound correlated with histomorphometry, Am. J. Cardiol., Vol. 78, pp. 1202–1209, 1996.

[45] Klingensmith, J., Shekhar, R., and Vince, D., Evaluation of three-dimensional segmentation algorithms for identification of luminal and medial-adventitial borders in intravascular ultrasound images, IEEE Trans. Med. Imaging, Vol. 19, No. 10, pp. 996–1011, 2000.

[46] McInerney, T. and Terzopoulos, D., Deformable models in medical images analysis: A survey, Med. Image Anal., Vol. 1, No. 2, pp. 91–108, 1996.

Chapter 3

Medical Image Segmentation: Methods and Applications in Functional Imaging

Koon-Pong Wong

3.1 Introduction

Detection, localization, diagnosis, staging, and monitoring treatment responses are the most important aspects and crucial procedures in diagnostic medicine and clinical oncology. Early detection and localization of the diseases and accurate disease staging can improve the survival and change management in patients prior to planned surgery or therapy. Therefore, current medical practice has been directed toward early but efficient localization and staging of diseases, while ensuring that patients would receive the most effective treatment.

Current disease management is based on the international standard of cancer staging using TNM classification, viz. size, location, and degree of invasion of the primary tumor (T), status of regional lymph node (N), and whether there is any distant metastasis (M). Over the century, histopathology retains its main role as the primary means to characterization of suspicious lesions and confirmation of malignancy. However, the pathologic interpretation is heavily dependent on the experience of the pathologist, and sampling errors may mean that there are insufficient amounts of tissue in the specimens, or the excised tissue does not accurately reflect tumor aggressivity. In addition, some lesions may return nondiagnostic information from the specimens, or they are difficult or too

[1]Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

dangerous to biopsy. As a result, more invasive and unpleasant diagnostic procedures are sought.

The last few decades of the twentieth century have witnessed significant advances in medical imaging and computer-aided medical image analysis. The revolutionary capabilities of new multidimensional medical imaging modalities and computing power have opened a new window for medical research and clinical diagnosis. Medical imaging modalities are used to acquire the data from which qualitative and quantitative information of the underlying pathophysiological basis of diseases are extracted for visualization and characterization, thus helping the clinicians to accurately formulate the most effective therapy for the patients by integrating the information with those obtained from some other possibly morbid and invasive diagnostic procedures. It is important to realize that medical imaging is a tool that is complementary to but not compete with the conventional diagnostic methods. Indeed, medical imaging provides additional information about the disease that is not available with the conventional diagnostic methods, and paves a way to improve our understanding of disease processes from different angles.

Modern medical imaging modalities can be broadly classified into two major categories: *structural* and *functional*. Examples of major *structural imaging* modalities include X-ray computed tomography (CT), magnetic resonance imaging (MRI), echocardiography, mammography, and ultrasonography. These imaging modalities allow us to visualize anatomic structure and pathology of internal organs. In contrast, *functional imaging* refers to a set of imaging techniques that are able to derive images reflecting biochemical, electrical, mechanical, or physiological properties of the living systems. Major functional imaging modalities include positron emission tomography (PET), single-photon emission computed tomography (SPECT), fluorescence imaging, and dynamic magnetic resonance imaging such as functional magnetic resonance imaging (fMRI) and magnetic resonance spectroscopy (MRS). Fundamentally, all these imaging techniques deal with reconstructing a three-dimensional image from a series of two-dimensional images (projections) taken at various angles around the body. In CT, the X-ray attenuation coefficient within the body is reconstructed, while in PET and SPECT our interest is in reconstructing the time-varying distribution of a labeled compound in the body in absolute units of radioactivity concentration.

Despite the differences between the actual physical measurements among different imaging modalities, the goal of acquiring the images in clinical environment is virtually the same—to extract the patient-specific clinical information

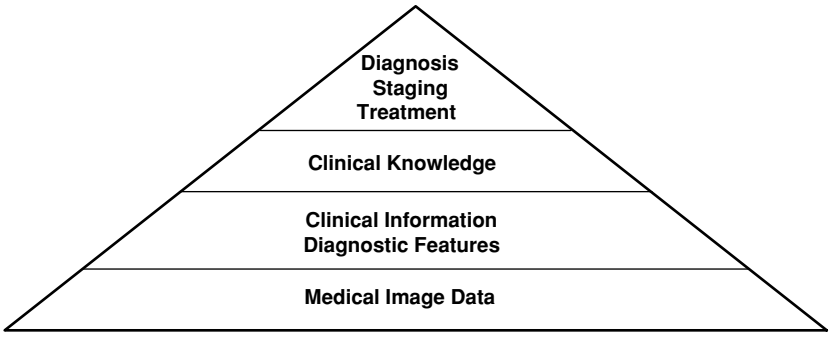


Figure 3.1: The steps and the ultimate goal of medical image analysis in a clinical environment.

and their diagnostic features embedded within the multidimensional image data that can guide and monitor interventions after the disease has been detected and localized, and ultimately leading to knowledge for clinical diagnosis, staging, and treatment of disease. These processes can be represented diagrammatically as a pyramid, as illustrated in Fig. 3.1. Starting from the bottom level of the pyramid is the medical image data obtained from a specific imaging modality, the ultimate goal (the top level of the pyramid) is to make use of the extracted information to form a set of clinical knowledge that can lead to clinical diagnosis and treatment of a specific disease. Now the question is how to reach the goal. It is obvious that the goal of the imaging study is very clear, but the solution is not. At each level of the pyramid, specific techniques are required to process the data, extract the information, label, and represent the information in a high level of abstraction for knowledge mining or to form clinical knowledge from which medical diagnosis and decision can be made. Huge amounts of multidimensional datasets, ranging from a few megabytes to several gigabytes, remain a formidable barrier to our capability in manipulating, visualizing, understanding, and analyzing the data. Effective management, processing, visualization, and analysis of these datasets cannot be accomplished without high-performance computing infrastructure composed of high-speed processors, storage, network, image display unit, as well as software programs. Recent advances in computing technology such as development of application-specific parallel processing architecture and dedicated image processing hardware have partially resolved most of the limiting factors. Yet, extraction of useful information and features from the multidimensional data is still a formidable task that requires specialized and sophisticated techniques. Development and implementation of these techniques requires

thorough understanding of the underlying problems and knowledge about the acquired data, for instance, the nature of the data, the goal of the study, and the scientific or medical interest, etc. Different assumptions about the data and different goals of the study will lead to the use of different methodologies. Therefore, continuing advances in exploitation and development of new conceptual approaches for effective extraction of all information and features contained in different types of multidimensional images are of increasingly importance in this regard.

Image segmentation plays a crucial role in extraction of useful information and attributes from images for all medical imaging applications. It is one of the important steps leading to image understanding, analysis, and interpretation. The principal goal of image segmentation is to partition an image into regions (or classes) that are homogeneous with respect to one or more characteristics or features under certain criteria [1]. Each of the regions can be separately processed for information extraction. The most obvious application of segmentation in medical imaging is anatomical localization, or in a generic term, region of interest delineation whose main aim is to outline anatomic structures and (pathologic) regions that are "of interest." Segmentation can be accomplished by identifying all pixels or voxels that belong to the same structure/region or based on some other attributes associated with each pixel or voxel. Image segmentation is not only important for feature extraction and visualization but also for image measurements and compression. It has found widespread applications in medical science, for example, localization of tumors and microcalcifications, delineation of blood cells, surgical planning, atlas matching, image registration, tissue classification, and tumor volume estimation [2–7], to name just a few.

Owing to issues such as poor spatial resolution, ill-defined boundaries, measurement noise, variability of object shapes, and some other acquisition artifacts in the acquired data, image segmentation still remains a difficult task. Segmentation of data obtained with functional imaging modalities is far more difficult than that of anatomical/structural imaging modalities, mainly because of the increased data dimensionality and the physical limitations of the imaging techniques. Notwithstanding these issues, there have been some significant progresses in this area, partly due to continuing advances in instrumentation and computer technology. It is in this context that an overview of the technical aspects and methodologies of image segmentation will be presented. As image

segmentation is a broad field and because the goal of segmentation varies according to the aim of the study and the type of the image data, it is impossible to develop only one standard method that suits all imaging applications. This chapter focuses on the segmentation of data obtained from functional imaging modalities such as PET, SPECT, and fMRI. In particular, segmentation based on cluster analysis, which has great potential in classification of functional imaging data, will be discussed in great detail. Techniques for segmentation of data obtained with structural imaging modalities have been covered in depth by other chapters of this book, and therefore, they will only be described briefly in this chapter for the purpose of completeness.

3.2 Manual Versus Automated Segmentation

As mentioned at the beginning of this chapter, detection, localization, diagnosis. staging, and monitoring treatment responses are crucial procedures in clinical medicine and oncology. Early detection and localization of the diseases and accurate disease staging could lead to changes in patient management that will impact on health outcomes. Noninvasive functional imaging is playing a key role in these issues. Accurate quantification of regional physiology depends on accurate delineation (or segmentation) of the structure or region of interest (ROI) in the images. The fundamental roles of ROI are to (1) permit quantitation, (2) reduce the dataset by focusing the quantitative analysis on the extracted regions that are of interest, and (3) establish structural correspondences for the physiological data sampled within the regions.

The most straightforward segmentation approach is to outline the ROIs manually. If certain areas in the images are of interest, the underlying tissue time–activity curve (TAC) can be extracted by putting ROIs manually around those areas. Approaches based on published anatomic atlases are also used to define ROIs. The average counts sampled over voxels in the region at different sampling intervals are then computed to compose the TAC for that region. The extracted tissue TACs are then used for subsequent kinetic analysis (Chapter 2 of *Handbook of Biomedical Image: Segmentation*, *Volume I*).

In practice, selection of ROI is tedious and time-consuming because the operator has to go through the dataset slice by slice (or even frame by frame) and choose the most representative ones from which 10–40 regions are carefully

delineated for each imaging study. Needless to say, manual ROI delineation is also operator dependent and the selected regions are subject to large intra- and interrater variability [8, 9]. Because of scatter and partial volume effects (PVEs) [10], the position, size, and shape of the ROI need careful consideration. Quantitative measurement inaccuracies exhibited by small positional differences are expected to be more pronounced for ROI delineation in the brain, which is a very heterogeneous organ and contains many small structures of irregular shape that lie adjacent to other small structures of markedly differing physiology [11]. Small positional differences can also confound the model fitting results [12, 13]. To minimize errors due to PVEs, the size of the ROI should be chosen as small as possible, but the trade-off is the increase in noise levels within the ROI, which maybe more susceptible to anatomical imprecision. On the other hand, a larger region offers a better signal-to-noise ratio but changes that occurred only within a small portion of the region maybe obscured, and the extracted TAC does not represent the temporal behavior of the ROI but a mixture of activities with adjacent overlapping tissue structures. Likewise, an irregular ROI that conforms to the shape of the structure/region where abnormality has occurred will be able to detect this change with much higher sensitivity than any other geometrically regular ROI that may not conform well. In addition, manual ROI delineation requires software tools with sophisticated graphical user interfaces to facilitate drawing ROIs and image display. Methodologies that can permit semiautomated or ideally, fully automated ROI segmentation will present obvious advantages over the manual ROI delineation.

Semiautomated or fully automated segmentation in anatomical imaging such as CT and MR is very successful, especially in the brain, as there are many well-developed schemes proposed in the literature (see surveys in [14]). This may be because these imaging modalities provide very high resolution images in which tiny structures are visible even in the presence of noise, and that four general tissue classes, gray matter, white matter, cerebrospinal fluid (CSF), and extracranial tissues such as fat, skin, and muscles, can be easily classified with different contrast measures. For instance, the T1- and T2-weighted MR images provide good contrast between gray matter and CSF, while T1 and proton density (PD) weighted MR images provide good contrast between gray matter and white matter. In contrast to CT and MRI, PET and SPECT images lack the ability to yield accurate anatomical information. The segmentation task is further complicated by poor spatial resolution and counting statistics, and patient motion

during scanning. Therefore, segmentation in PET and SPECT does not attract much interest over the last two decades, even though there has been remarkable progress in image segmentation during the same period of time. It still remains a normal practice to define ROIs manually.

Although the rationale for applying automatic segmentation to dynamic PET and SPECT images is questionable due to the above difficulties, the application of automatic segmentation as an alternative to manual ROI delineation has attracted interest recently with the improved spatial resolution of PET and SPECT systems. Automatic segmentation has advantages in that the subjectivity can be reduced and that there is saving in time for manual ROI delineation. Therefore, it may provide more consistent and reproducible results as less human intervention is involved, while the overall time for data analysis can be shortened and thereby the efficiency can be improved, which is particularly important in busy clinical settings.

3.3 Optimal Criteria for Functional Image Segmentation

Medical image segmentation is a very complicated process and the degree of complexity varies under different situations. Based on the results of a survey conducted among all centers performing emission tomographic studies and a series of international workshops to assess the goals and obstacles of data acquisition and analysis from emission tomography, Mazziotta *et al.* [15, 16] proposed a series of optimal criteria to standardize and optimize PET data acquisition and analysis:

- Reproducible
- Accurate
- Independent of tracer employed
- Independent of instrument spatial resolution
- Independent of ancillary imaging techniques
- Minimizes subjectivity and investigator bias

- Fixed assumptions about normal anatomy not required
- Acceptable to subjects' level of tolerance
- Performs well in serial studies of the same patient and individual study of separate patients in a population
- Capable of evolving toward greater accuracy as information and instruments improve
- Reasonable in cost
- Equally applicable in both clinical and research settings
- Time efficient for both data acquisition and analysis

These criteria are not specific to the functional analysis of the brain, and they are equally applicable to other organs and imaging applications upon minor modifications, in spite of the fundamentally differences between imaging modalities.

3.4 Segmentation Techniques

A large number of segmentation techniques have been proposed and implemented (see [14, 17–19]) but there is still no *gold standard* approach that satisfies all of the aforementioned criteria. In general, segmentation techniques can be divided into four major classes:

- Thresholding
- Edge-based segmentation
- Region-based segmentation
- Pixel classification

These techniques are commonly employed in two-dimensional image segmentation [1, 17–21]. A brief review of these techniques will be given in this section. More advanced techniques such as model-based approaches, multimodal approaches, and multivariate approaches, and their applications will be introduced and discussed latter in this chapter.

3.4.1 Thresholding

Semiautomatic methods can partially remove the subjectivity in defining ROIs by human operators. The most commonly used method is by means of thresholding because of its simplicity in implementation and intuitive properties. In this technique, a predefined value (threshold) is selected and an image is divided into groups of pixels having values greater than or equal to the threshold and groups of pixels with values less than the threshold. The most intuitive approach is global thresholding, which is best suited for bimodal image. When only a single threshold is selected for a given image, the thresholding is *global* to the entire image. For example, let $f(x, y)$ be an image with maximum pixel value $I_{\max}$, and suppose Γ denotes the percent threshold of the maximum pixel value above which the pixels will be selected, then pixels with value ρ given by

$$\frac{\Gamma}{100} I_{\max} \leq \rho \leq I_{\max} \tag{3.1}$$

can be grouped and a binary image $g(x, y)$ is formed:

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq \rho \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

in which pixels with value of 1 correspond to the ROI, while pixels with value 0 correspond to the background.

Global thresholding is simple and computationally fast. It performs well if the images contain objects with homogeneous intensity or the contrast between the objects and the background is high. However, it may not lead itself fully automated and may fail when two or more tissue structures have overlapping intensity levels. The accuracy of the ROI is also questionable because it is separated from the data based on a single threshold value which may be subject to very large statistical fluctuations. With the increasing number of regions or noise levels, or when the contrast of the image is low, threshold selection will become more difficult.

Apart from global thresholding, there are several thresholding methods which can be classified as local thresholding and dynamic thresholding. These techniques maybe useful when a thresholding value cannot be determined from a histogram for the entire image or a single threshold cannot give good segmentation results. Local threshold can be determined by using the local statistical properties such as the mean value of the local intensity distribution or some

other statistics such as mean of the maximum or minimum values [21] or local gradient [22], or by splitting the image into subimages and calculating threshold values for the individual sub-images [23]. Some variants of the above two methods can be found in Refs. [17, 18].

3.4.2 Edge-Based Segmentation

Edge-based segmentation approaches have two major components: (1) edge detection and (2) edge linking/following to determine the edges and the regions. Loosely speaking, an edge is a collection of connected pixels that lie on the boundary between two homogeneous regions having different intensities. Therefore, edges can be defined as abrupt changes in pixel intensity that can be reflected by the gradient information. A number of edge detectors have been defined based on the first-order or second-order gradient information of the image [1, 20]. For a given image $f(x, y)$, the gradient computed at location (x, y) is defined as a vector:

$$\nabla \mathbf{f} = \begin{bmatrix} \delta f_x \\ \delta f_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix} \tag{3.3}$$

where δf_x and δf_y are gradients computed along x and y directions. The gradient vector points in the direction of maximum rate of change of f at (x, y). The magnitude and the direction of the gradient vector are given by

$$|\nabla \mathbf{f}| = \sqrt{(\delta f_x)^2 + (\delta f_y)^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \tag{3.4}$$

and

$$\theta = \tan^{-1}\left(\frac{\delta f_y}{\delta f_x}\right) \tag{3.5}$$

where the angle θ is measured with respect to the x axis.

In order to obtain the gradient of an image, computation of partial derivatives δf_x and δf_y at every pixel location is required. Because the images have been digitized, it is not possible to compute δf_x and δf_y using differentiation and numerical approximation of the gradient by finite difference is used instead [20]. Implementation of edge detection can be accomplished by convolving the original image with a *mask* (also called *window* or *kernel*) that runs through the

entire image. A mask is typically a 2×2 or a 3×3 matrix. For example, Roberts edge operator has two 2×2 masks:

$$\delta f_x = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \delta f_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and Sobel edge operator has a pair of 3×3 masks:

$$\delta f_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad \delta f_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Detailed discussion on other edge operators such as Canny, Kirsch, Prewitt, and Robinson can be found elsewhere [1, 20].

An edge magnitude image can be formed by combining the gradient components δf_x and δf_y at every pixel location using Eq. (3.4). As the computational burden required by square and square roots in Eq. (3.4) is very high, an approximation with absolute values is frequently used instead:

$$|\nabla \mathbf{f}| \approx |\delta f_x| + |\delta f_y| \tag{3.6}$$

After the edge magnitude image has been formed, a thresholding operation is then performed to determine where the edges are.

The first-order derivatives of the image $f(x, y)$ have local minima and maxima at the edges because of the large intensity variations. Accordingly, the second-order derivatives have zero crossings at the edges, which can also be used for edge detection and the Laplacian is frequently employed in practice. The Laplacian (∇^2) of a two-dimensional function $f(x, y)$ is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{3.7}$$

There are several ways to realize the Laplacian operator in discrete-time domain. For a 3×3 region, the following two realizations are commonly used:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

It should be noted that all gradient-based edge detection methods (including the Laplacian) are very sensitive to noise because differentiation is a high pass

operation that tends to enhance image noise. In some applications, it is possible to improve the results obtained with these methods by smoothing the image prior to edge detection. For example, a smoothing filter can be applied to an image before computing the Laplacian. Marr and Hildreth [24] proposed smoothing the image with a Gaussian filter followed by the Laplacian operation to determine edge information and this operation is called Laplacian of Gaussian, which is defined as

$$\begin{aligned} h(x, y) &= \nabla^2[g(x, y) \otimes f(x, y)] \\ &= \nabla^2[g(x, y)] \otimes f(x, y) \end{aligned} \tag{3.8}$$

where $f(x, y)$ is the original image, $\otimes$ is the convolution operator, $g(x, y)$ is a Gaussian function, and $\nabla^2[g(x, y)]$ is the Laplacian of Gaussian function that is used for spatial smoothing of the original image. Edges can be determined from the resultant image, $h(x, y)$, by simply thresholding it for zero value to detect zero crossing. Equation (3.8) represents a generic operation of taking the Laplacian on the spatial smoothing filter, $g(x, y)$, which can be replaced by other filter function (e.g., directional low-pass filter [25]) to improve the performance of edge detection in a specific application. Faber *et al.* [26] applied the Laplacian edge detection technique to segment scintigraphic images and the results were promising.

In practice, edge detection techniques produce only a series of edges for the structures/areas of interest. It is not uncommon that the edge pixels do not characterize an edge and that the edges do not enclose the ROI completely because of noise and some other acquisition artifacts that caused spurious discontinuities of edges. Therefore, the second component of edge-based segmentation techniques is to track and link the edge pixels to form meaningful edges or closed regions in the edge image obtained by one of the edge detection techniques. One of the simplest approaches for linking edge pixels is to analyze local characteristics of pixels within a small block of pixels (e.g., 3×3 or 5×5) for the entire edge image and linked all edge pixels that are similar to each others according to some predefined criteria. The Hough transform [27] can also be applied to detect straight lines and parametric curves as well as arbitrarily shaped objects in the edge image. It was shown by Deans [28] that the Hough transform is a special case of the Radon transform for image projection and reconstruction [29]. Thorough discussion and comparison of different varieties of the Hough transform and their generalizations are considered beyond the

scope of this chapter and they can be found in Refs. [30, 31]. There are several more powerful edge tracking/linking techniques such as graph searching [32,33] and dynamic programming [34, 35] that perform well in the presence of noise. As might be expected, these paradigms are considerably more complicated and computationally expensive than the methods discussed so far.

3.4.3 Region-Based Segmentation

Region-based segmentation approaches examine pixels in an image and form disjoint regions by merging neighborhood pixels with homogeneity properties based on a predefined similarity criterion. Suppose that I represents an image that is segmented into N regions, each of which is denoted as R_i where $i = 1, 2, \ldots, N$, the regions must satisfy the following properties:

$$I = \bigcup_{i=1}^{N} R_i \tag{3.9}$$

$$R_i \cap R_j = 0 \;\; \forall i, j = 1, 2, \ldots, N; \;\; i \neq j \tag{3.10}$$

$$\mathcal{L}(R_i) = \text{TRUE} \quad \text{for } i = 1, 2, \ldots, N \tag{3.11}$$

$$\mathcal{L}(R_i \cup R_j) = \text{FALSE} \quad \forall i, j = 1, 2, \ldots, N; \;\; i \neq j \tag{3.12}$$

where $\mathcal{L}(\cdot)$ is a logical predicate. The original image can be exactly assembled by putting all regions together (Eq. 3.9) and there should be no overlapping between any two regions R_i and R_j for $i \neq j$ (Eq. 3.10). The logical predicate $\mathcal{L}(\cdot)$ contains a set of rules (usually a set of homogeneity criteria) that must be satisfied by all pixels within a given region (Eq. 3.11), and it fails in the union of two regions since merging two distinct regions will result in an inhomogeneous region (Eq. 3.12).

The simplest region-based segmentation technique is the *region growing*, which is used to extract a connected region of similar pixels from an image [36]. The region growing algorithm requires a similarity measure that determines the inclusion of pixels in the region and a stopping criterion that terminates the growth of the region. Typically, it starts with a pixel (or a collection of pixels) called *seed* that belongs to the target ROI. The seed can be chosen by the operator or determined by an automatic seed finding algorithm. The neighborhood of each seed is then inspected and pixels similar enough to the seed are added to the corresponding region where the seed is, and therefore, the region is growing and its shape is also changing. The growing process is repeated until no pixel

can be added to any region. It is possible that some pixels may remain unlabeled when the growing process stops.

Hebert *et al.* [37] investigated the use of region growing to automated delineation of the blood pool with computer simulations and applied the method to three gated SPECT studies using Tc-99m pertechnetate, and the results were promising. Kim *et al.* [38] also investigated an integrated approach of region growing and cluster analysis (to be described later) to segment a dynamic [^{18}F]fluorodeoxyglucose (FDG) PET dataset. Although qualitatively reasonable segmentation results were obtained, much more work is needed to overcome the difficulties in the formation of odd segments possibly due to spillover region boundaries, and evaluate the quantitative accuracy of the segmentation results using kinetic parameter estimation.

Region splitting methods take an opposite strategy to the region growing. These methods start from the entire image and examine the homogeneity criteria. If the criteria do not meet, the image (or subimage) is split into two or more subimages. The region splitting process continues until all subimages meet the homogeneity criteria. Region splitting can be implemented by quadtree partitioning. The image is partitioned into four subimages that are represented by nodes in a quadtree, which is a data structure used for efficient storage and representation. The partition procedure is applied recursively on each subimage until each and all of the subimages meet the homogeneity criteria.

The major drawback of region splitting is that the final image may contain adjacent regions R_i and R_j, which are homogeneous, i.e. $\mathcal{L}(R_i \cup R_j) = \text{TRUE}$, and ideally this region should be merged. This leads to another technique called *split-and-merge*, which includes a merging step in the splitting stage, where an inhomogeneous region is split until homogeneous regions are formed. A newly created homogeneous region is checked against its neighboring regions and merged with one or more of these regions if they possess identical properties. However, this strategy does not necessarily produce quadtree partitioning of the image. If quadtree partitioning is used, an additional step may be added to merge adjacent regions (nodes) that meet the homogeneity criterion.

3.4.4 Pixel Classification

Recall that the key step of thresholding techniques described in section 3.4.1 is the choice of thresholds that is determined either manually or in a

semiautomatic manner based on the local statistics such as mean, maximum, or minimum of the given image (or subimages). The basic concept of threshold selection can be generalized, leading to a data-driven paradigm, which determines the threshold automatically based on clustering techniques or artificial neural networks.

Pixel classification methods that use histogram statistics to define single or multiple thresholds to classify an image can be regarded as a generalization of thresholding techniques. It is particularly useful when the pixels have multiple features, which can be expressed in terms of a vector in multidimensional feature space. For instance, the feature vector may consist of gray level, local texture, and color components for each pixel in the image. In the case of single-channel (or single-frame) image, pixel classification is typically based on gray level and image segmentation can be performed in a one-dimensional feature space. Segmentation can be performed in multidimensional feature space through clustering of all features of interest for multichannel (multiple-frame) images or multispectral (multimodality) images.

Clustering, or cluster analysis, has been widely applied in anthropology, archaeology, psychiatry, and zoology, etc, for many years. An example of clustering is the taxonomy of animals and plants whose names have to be the same between different people for effective communication, although it is not necessary that the naming scheme is the best [39]. Clustering is the process of grouping of similar objects into a single cluster, while objects with dissimilar features are grouped into different clusters based on some similarity criteria. The similarity is quantified in terms of an appropriate distance measure. An obvious measure of the similarity is the distance between two vectors in the feature space which can be expressed in terms of L_p norm as

$$d\{\mathbf{x}_i, \mathbf{x}_j\} = \left(\sum_{k=1}^{n} \| \mathbf{x}_i - \mathbf{x}_j \|^p \right)^{\frac{1}{p}} \tag{3.13}$$

where $\mathbf{x}_i \in \mathbb{R}^n$ and $\mathbf{x}_j \in \mathbb{R}^n$ are the two vectors in the feature space. It can be seen that the above measure corresponds to Euclidean distance when $p = 2$ and Mahalanobis distance when $p = 1$. Another commonly used distance measure is the normalized inner product between two vectors given by

$$d\{\mathbf{x}_i, \mathbf{x}_j\} = \frac{\mathbf{x}_i^T \mathbf{x}_j}{\|\mathbf{x}_i\| \cdot \|\mathbf{x}_j\|} \tag{3.14}$$

where T denotes the transpose operation. The above distance measure is simply the angle between vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ in the feature space.

Each cluster is represented by its centroid (or mean) and variance, which indicates the compactness of the objects within the cluster, and the formation of clusters is optimized according to a cost function that typically takes the similarity within individual cluster and dissimilarity between clusters into account. There are many clustering techniques proposed in the literature (see Ref. [39]). The most famous clustering techniques are K-means [40], fuzzy c-means [41], ISODATA [42], hierarchical clustering with average linkage method [43], and Gaussian mixture approach [44].

As we will see later in this chapter, the idea of pixel classification in two-dimensional image segmentation using clustering techniques can be extended to multidimensional domain where the images convey not only spatial information of the imaged structures but also their temporal variations, for which clustering plays a pivotal role in identification of different temporal kinetics present in the data, extraction of blood and tissue TACs, ROI delineation, localization of abnormality, kinetic modeling, characterization of tissue kinetics, smoothing, and fast generation of parametric images.

3.5 Advanced Segmentation Techniques

Functional imaging with PET, SPECT, and/or dynamic MRI provides *in vivo* quantitative measurements of physiologic parameters of biochemical pathways and physiology in a noninvasive manner. A critical component is the extraction of physiological data, which requires accurate localization/segmentation of the appropriate ROIs. A common approach is to identify the anatomic structures by placing ROIs directly on the functional images, and the underlying tissue TACs are then extracted for subsequent analysis. This ROI analysis approach, although widely used in clinical and research settings, is operator-dependent and thus prone to reproducibility errors and it is also time-consuming. In addition, this approach is problematic when applied to small structures because of the PVEs due to finite spatial resolution of the imaging devices.

Methods discussed so far can be applied to almost all kinds of image segmentation problem because they do not require any model (i.e. model-free) that guides or constrains the segmentation process. However, segmenting structures

of interest from functional images is difficult because of the imprecise anatomical information, the complexity and variability of anatomy shapes and sizes within and across individuals, and acquisition artifact, such as spatial aliasing, and insufficient temporal sampling, noise, and organ/patient movements. All these factors can hamper the boundary detection process and cause discontinuous or indistinguishable boundaries. Model-free approaches usually generate ambiguous segmentation results under these circumstances, and considerable amounts of human intervention are needed to resolve the ambiguity in segmentation. In this section, some advanced segmentation approaches are introduced, including

- *model-based* segmentation techniques that use analytical models to describe the shape of the underlying ROI,
- *multimodal techniques* that integrate information available from different imaging modalities for segmentation, or the image measurements are transformed and mapped to a standard template, and
- *multivariate approaches* are data-driven techniques in which the structures are identified and extracted based on the temporal information present in the dynamic images.

3.5.1 Model-Based Segmentation

Incorporation of *a priori* knowledge of the object such as shape, location, and orientation using *deformable models* (also known as *active contour models*) is one of the possible solutions to constrain the segmentation of organ structures. The term *deformable models* was coined by Terzopoulos and his collaborators [45,46] in the 1980s, but the idea of using a deformable template for feature extraction dated back to the work of Fischler and Elschlager on spring-loaded templates [47] and the work of Widrow on rubber mask technique [48] in the early 1970s. Deformable models are analytically or parametrically defined curves or surfaces that move under the influence of forces, which have two components: *internal forces* and *external forces*. The internal forces are used to assure the smoothness of the model during deformation process and the external forces are defined to push/pull the model toward the boundaries of the structure. Parametric representations of the models allow accurate and compact description

of the object shape, while the continuity, connectivity, and smoothness of the models can compensate for the irregularities and noise in the object boundaries. Model-based approaches treat the problem of finding object boundaries as an optimization problem of searching the best fit for the image data to the model. In the case of boundary finding via optimization in image space, a fairly extensive review on various deformable model methods can be found in Ref. [49].

Mykkänen *et al.* [50] investigated automatic delineation of brain structures in FDG-PET images using generalized snakes with promising results. Chiao *et al.* [51] proposed using model-based approach for segmenting dynamic cardiac PET or SPECT data. The object model consists of two parts: a heart and the rest of the body. The heart is geometrically modeled using a polygonal model [52] and the myocardial boundaries are parameterized by the endocardial radii and a set of angular thicknesses. Kinetic parameters in the compartment model and the endocardial and epicardial radii are estimated by maximizing a joint log-likelihood function using nonlinear parameter estimation. Tissue and blood TACs are extracted simultaneously with estimated kinetic parameters. Chiao *et al.* [51] proposed that some forms of regularization can be applied, including auxiliary myocardial boundary measurements obtained by MRI or CT and registering the auxiliary measurements with the emission tomographic data, if the kinetic parameter estimation failed.

3.5.2 Multimodal Techniques

Comparisons of datasets obtained from individual subjects between imaging modalities are very important for the evaluation of the normal physiologic responses of the anatomic structure or the pathophysiological changes that accompany disease states. Likewise, it is also critical to compare data between individuals both within and across different imaging modalities. Unfortunately, many structures of interest, particularly in the brain, are often smaller than the spatial resolution of the imaging devices and corrections aided by anatomical imaging modalities such as CT and MR are often required [53, 54].

Anatomic structures, particularly those in the brain, can also be identified using a standardized reference coordinate system or functional image data can be fitted to a standard anatomical atlas (e.g., Talairach space) with the aid of anatomical landmarks or contours [55–58]. This idea is somewhat similar to the model-based approaches where analytically or parametrically defined models

are used to segment the organ boundaries. The difference lies in the definition of the model, which is described by a computerized anatomy atlas or a stereotaxic coordinate system—a reference that the functional images are mapped onto by either linear or nonlinear transformation. A number of transformation techniques have been developed for this process [59]. The ROIs defined on the template are then available to the functional image data.

Similarly, functional (PET and SPECT) images and structural (CT and MR) images obtained from individual subjects can be fused (coregistered), allowing precise anatomical localization of activity on the functional images [60, 61]. Precise alignment between the anatomic/template and PET images is necessary for these methods. Importantly, methods that use registration to a standard coordinate system are problematic when patients with pathological processes (e.g., tumors, infarction, and atrophy) are studied.

3.5.3 Multivariate Segmentation

The main aim of dynamic imaging is to study the physiology (function) of the organ *in vivo*. Typically the image sequence has constant morphologic structures of the imaged organs but the regional voxel intensity varies from one frame to another, depending on the local tissue response to the administered contrast agent or radiopharmaceutical. In the past, analysis of such dynamic images involved only visual analysis of differences between the early and delayed images from which qualitative information about the organ, for instance, regional myocardial blood flow and distribution volume are obtained. However, the sequence of dynamic images also contain spatially varying quantitative information about the organ which is difficult to extract solely based on visual analysis. This led to the method of parametric imaging where dynamic curves in the image sequence are fit to a mathematical model on a pixel-wise basis. Parametric images whose pixels define individual kinetic parameters or physiologic parameters that describe the complex biochemical pathways and physiologic/pharmacokinetic processes occurred within the tissue/organ can then be constructed. This approach is categorized as *model-led* technique that utilizes knowledge and *a priori* assumptions of the processes under investigation, and represents the kinetics of the measured data by an analytical (or parametric) model.

At the opposite end of the spectrum of model-led techniques are *data-driven* techniques, which are based on the framework of multivariate data analysis.

These paradigms minimize the *a priori* assumptions of the underlying processes whose characteristics are interrogated from the measured data, independent of any kinetic model. Multivariate approaches have been explored and successfully applied in a number of functional imaging studies. The aims borne in mind when applying these approaches are to (1) classify structures present in the images, (2) extract information from the images, (3) reduce the dimensionality of data, (4) visualize the data, and (5) model the data, all of which are crucial in data analysis, medical research, data reduction, and treatment planning. In general, the underlying structures are identified and extracted based on the temporal information present in the sequence of the medical images. The implicit assumptions for the validity of applying these approaches are that the statistical noise present in the images is uncorrelated (independent) between different frames and that there is a high degree of correlation (similarity) between tissue TACs if they are originated from similar structures. In this section, we focus our attention on four techniques among many different multivariate analysis approaches and their applications in dynamic, functional imaging are discussed.

3.5.3.1 Similarity Mapping

In this section, we introduce an intuitive temporal segmentation technique called *similarity mapping* (or *correlation mapping*), which was proposed by Rogowska [62]. This approach identifies regions according to their temporal similarity or dissimilarity with respect to a dynamic curve obtained from a reference region. Consider a sequence of N spatially registered time-varying images $\mathbf{X}$ of size $M \times N$, with M being the number of pixels in one image and N the number of frames. Then each row of $\mathbf{X}$ represents a pixel vector, i.e., a time-intensity curve as stated in Rogowska's paper [62] (also called a *dixel* [63] or a tissue TAC in PET/SPECT or fMRI studies—it is just a matter of nomenclature!) which is a time series

$$\mathbf{x}_i = [X_i(t_1), X_i(t_2), \ldots, X_i(t_N)]^T \tag{3.15}$$

where t_j ($j = 1, 2, \ldots, N$) represents the time instant at which the jth frame is acquired, $X_i(t_j)$ is the pixel value of the ith element evaluated in the jth frame of $\mathbf{X}$, for $j = 1, 2, \ldots, N$, and T denotes the transpose operation.

Similar to the pixel classification technique described earlier in section 3.4.4, some quantitative index is necessary to measure the similarity between

time-intensity curves for different pixels or the mean of the pixel values averaged over a selected ROI. Suppose a ROI is drawn on a reference region in the dynamic sequence of images and its time course is extracted

$$\mathbf{r} = [r(t_1), r(t_2), \ldots, r(t_N)]^T \tag{3.16}$$

The similarity between the reference time-intensity curve $\mathbf{r}$ and the time-intensity curves for all pixels can then be calculated. And a similarity map, which is a image where the value of each pixel shows the temporal similarity to the reference curve, can be constructed.

Since the time instants do not affect the computation of cross correlation between two time-intensity curves as pixel intensity values in one frame are measured at the same time, $\mathbf{x}_i$ in Eq. (3.15) and $\mathbf{r}$ in Eq. (3.16) can be rewritten in a time-independent form as

$$\mathbf{x}_i = [X_{i,1}, X_{i,2}, \ldots, X_{i,N}]^T \tag{3.17}$$

where $X_{i,j} \equiv X_i(t_j)$ is the pixel value of the ith element evaluated in the jth frame of $\mathbf{X}$, and

$$\mathbf{r} = [r_1, r_2, \ldots, r_N]^T \tag{3.18}$$

whose mean intensity value is given by

$$\overline{r} = \frac{1}{N}\sum_{j=1}^{N} r_j \tag{3.19}$$

The similarity map $\mathbf{R}$ based on normalized cross correlation can be defined for each pixel i as

$$R_i = \frac{\sum_{j=1}^{N}\left(X_{i,j} - \overline{X}_i\right)\left(r_j - \overline{r}\right)}{\sqrt{\sum_{j=1}^{N}\left(X_{i,j} - \overline{X}_i\right)^2 \sum_{j=1}^{N}\left(r_j - \overline{r}\right)^2}} \tag{3.20}$$

where

$$\overline{X}_i = \frac{1}{N}\sum_{j=1}^{N} X_{i,j} \tag{3.21}$$

is the mean value of the time sequence for pixel i. The normalized cross correlation has values in the range of -1 to $+1$. Regions of identical temporal variation have a coefficient of $+1$, with the exception that $\mathbf{x}_i$ or $\mathbf{r}$ are extracted from regions of constant pixel intensity (e.g. background). In this case, the denominator of Eq. (3.20) equals zero. Therefore, the following restrictions have to be

imposed on the computation of the normalized cross correlation:

$$\sum_{j=1}^{N} \left(X_{i,j} - \overline{X}_i\right)^2 \neq 0 \quad \text{and} \quad \sum_{j=1}^{N} \left(r_j - \overline{r}\right)^2 \neq 0 \tag{3.22}$$

Time-intensity curves similar to the reference curve will have high-correlation values and are bright in the similarity map, whereas those with low-correlation values are dark. Therefore, structures in the dynamic image sequence can be segmented from the similarity map based on their temporal changes rather than spatial similarities. It should be noted that cross-correlation does not depend on the absolute magnitude of the time-intensity curves. Regions whose time-intensity curves are differed from the reference curve **r** by an additive or by a multiplicative constant, will have a perfect positive correlation (+1).

By using different reference ROIs, a series of similarity maps containing different segmentation for regions that have similar or different temporal kinetics can be obtained. It was used to investigate local changes and segmentation of rabbit kidney on spatially aligned image sequences obtained from dynamic MR imaging of Gd-DTPA [64]. The similarity mapping technique has also been applied to brain activation studies to extract the activated regions and their temporal dynamics [65]. The same technique has also been used to segment the area of ischemia in the left coronary artery territory, lung tumor, and tentorial meningioma, and localize the focal ischemic region in brain [66].

3.5.3.2 Principal Component Analysis

Principal component analysis (PCA), also called *Karhunen–Loéve transform* or *Hotelling transform*, is probably the most famous method in multivariate data analysis [67]. It was developed independently by Pearson [68] in 1901 and Hotelling [69] in 1933. It has been widely applied in a number of scientific areas such as biology, chemistry, medicine, psychology, and the behavioral sciences. Given a set of multivariate data, PCA explains the variance–covariance structure by linearly transforming the (possibly) correlated variables into a smaller set of uncorrelated (orthogonal) variables called *principal components* (PCs). The first (highest order) component maximally accounts for the variation in the original data and each succeeding component maximally accounts for the remaining variation present in the original data. In other words, higher order components are important as they explain the major variation (also the feature) in

the data, whereas lower order components are unimportant as they mainly contain noise, which can be discarded without causing too much loss of information of the original data. Therefore, dimensionality reduction (or data compression) can be achieved using PCA technique. Separation of tissue types characterized by different features can also be accomplished by careful inspection of the PCs. This is because each PC contains only the representative feature that is specific to that PC and cannot be found elsewhere (theoretically) owing to orthogonality among PCs.

Let the dynamic sequence of images be represented by a matrix $\mathbf{X}$ that has M rows and N columns. Each column represents a time frame of image data and each row represents a pixel vector, i.e., a tissue TAC or a *dixel* [63], which is a time series $\mathbf{x}_i$ as in Eqs. (3.15) and (3.17). Note that there is no explicit assumption on the probability density of the measurements $\mathbf{x}_i$ as long as the first-order and second-order statistics are known or can be estimated from the available measurements. Each of $\mathbf{x}_i$ can be considered as a random process

$$\mathbf{x} = [x_1, x_2, \ldots, x_N]^T \tag{3.23}$$

If the measurements (or random variables) x_j are correlated, their major variations can be accurately approximated with less than N parameters using the PCA. The mean of $\mathbf{x}$ is given by

$$\overline{x} = E\{\mathbf{x}\} \tag{3.24}$$

and the covariance matrix of the same dataset is given by

$$\mathbf{C}_\mathbf{x} = E\{(\mathbf{x} - \overline{x})(\mathbf{x} - \overline{x})^T\} \tag{3.25}$$

which is a $N \times N$ symmetric matrix. The elements of $\mathbf{C}_\mathbf{x}$, denoted by c_{kl}, represent the covariances between the random variables x_k and x_l, whereas the element c_{kk} is the variance of the random variable x_k. If x_k and x_l are uncorrelated, their covariance would be zero, i.e., $c_{kl} = c_{lk} = 0$. The mean and the covariance matrix of a sample of random vectors $\mathbf{x}_i$ can be estimated from its sample mean and sample covariance matrix in a similar manner.

The orthogonal basis of the covariance matrix $\mathbf{C}_\mathbf{x}$ can be calculated by finding its eigenvalues and eigenvectors. It is well known from basic linear algebra that the eigenvectors $\mathbf{e}_k$ and the corresponding eigenvalues λ_k are the solutions of

the equation

$$\mathbf{C}_\mathbf{x}\mathbf{e}_k = \lambda_k \mathbf{e}_k \tag{3.26}$$

for $k = 1, 2, \ldots, N$ and $\lambda_k \neq 0$. There are several numerical methods to solve for λ_k and $\mathbf{e}_k$ in Eq. (3.26). One of the popular approaches is to make use of the symmetrical property of $\mathbf{C}_\mathbf{x}$ and solve for the eigenvalues and eigenvectors by means of Householder reduction followed by QL algorithm with implicit shifts [70, 71]. As $\mathbf{C}_\mathbf{x}$ is a real, symmetric matrix, an equivalent approach is to compute the singular value decomposition (SVD) of the matrix $\mathbf{C}_\mathbf{x}$ directly:

$$\mathbf{C}_\mathbf{x} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \tag{3.27}$$

where $\mathbf{U}$ is a $N \times N$ column-orthogonal matrix, $\mathbf{V}$ is a $N \times N$ orthogonal matrix that contains the eigenvectors, and $\mathbf{\Lambda}$ is a $N \times N$ diagonal matrix whose squared diagonal elements correspond to the eigenvalues. However, the difference between SVD and the eigen-decomposition should be noted, in particular, the eigen-decomposition of a real matrix might be complex, whereas the SVD of a real matrix is always real.

The ordering of the eigenvectors can be sorted in the order of descending eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$. In this way, an ordered orthogonal basis is formed, and the first eigenvector $\mathbf{e}_1$ (the one associated with λ_1) has the direction of the largest variance of the data (the first PC), and the second eigenvector $\mathbf{e}_2$ has the direction of the second largest variance of the data (the second PC), and so on. The PCs are obtained by projecting the multivariate random vectors onto the space spanned by the eigenvectors. Let $\mathbf{\Omega}$ be a matrix that stores the eigenvectors $\mathbf{e}_k$ as row vectors, then the PCs, $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$, can be calculated as

$$\mathbf{y} = \mathbf{\Omega}(\mathbf{x} - \overline{x}) \tag{3.28}$$

which defines a linear transformation for the random vector $\mathbf{x}$ through the orthogonal basis and $\overline{x}$ is calculated from Eq. (3.24). The kth PC of $\mathbf{x}$ is given by

$$y_k = \mathbf{e}_k^T(\mathbf{x} - \overline{x}) \tag{3.29}$$

which has zero mean. The PCs are also orthogonal (uncorrelated) to one another because

$$E\{y_k y_l\} = E\left\{\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\left(\mathbf{e}_l^T(\mathbf{x}-\overline{x})\right)\right\} = \mathbf{e}_k^T\mathbf{C}_\mathbf{x}\,\mathbf{e}_l = 0 \tag{3.30}$$

for $k > l$. The original random vector $\mathbf{x}$ can be reconstructed from $\mathbf{y}$ by

$$\mathbf{x} = \mathbf{\Omega}^T\mathbf{y} + \overline{x} \tag{3.31}$$

where $\mathbf{\Omega}^{-1} = \mathbf{\Omega}^T$ since $\mathbf{\Omega}$ is an orthogonal matrix.

The variances of the PCs can be computed as follows:

$$E\{y_k^2\} = E\left\{\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\left(\mathbf{e}_k^T(\mathbf{x}-\overline{x})\right)\right\} = \mathbf{e}_k^T\mathbf{C}_\mathbf{x}\,\mathbf{e}_k = \lambda_k \tag{3.32}$$

which indicates that the variances of the PCs are given by the eigenvalues of $\mathbf{C}_\mathbf{x}$. As the PCs have zero means, a very small eigenvalue (variance) λ_k implies that the value of the corresponding PC is also very small to contribute to the total variances present in the data. Since the eigenvalue sequence $\{\lambda_k\}$ is monotonically decreasing and typically the sequence drops rapidly, it is possible to determine a limit below which the eigenvalues (and PCs) can be discarded without causing significant error in reconstruction of the original dataset using only the retained PCs. Thus, data compression (or dimensionality reduction) can be achieved and this is an important application of PCA. Instead of using all eigenvectors of the covariance matrix $\mathbf{C}_\mathbf{x}$, the random vector $\mathbf{x}$ can be approximated by the highest few basis vectors of the orthogonal basis. Suppose that only the first K rows (eigenvectors) of $\mathbf{\Omega}$ are selected to form a $K \times N$ matrix, $\mathbf{\Omega}_K$, a similar transformation as in Eqs. (3.28) and (3.31) can be derived

$$\tilde{\mathbf{y}} = \mathbf{\Omega}_K(\mathbf{x} - \overline{x}) \tag{3.33}$$

and

$$\hat{\mathbf{x}} = \mathbf{\Omega}_K^T\mathbf{y} + \overline{x} \tag{3.34}$$

where $\tilde{\mathbf{y}}$ represents a truncated PC vector, which contains only the K highest PCs, and $\hat{\mathbf{x}}$ is an approximation of $\mathbf{x}$ with the K highest PCs. It can be shown that the mean square error (MSE) between $\hat{\mathbf{x}}$ and $\mathbf{x}$ is given by

$$E\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} = \sum_{k=K+1}^{N} \lambda_k \tag{3.35}$$

The practical issue here is the choice of K beyond which the PCs are insignificant. The gist of the problem lies in how "insignificant" is defined and how much error one could tolerate in using less number of PCs to approximate the original data. Sometimes, a small number of PCs are sufficient to give an accurate approximation to the observed data. A commonly used strategy is to plot the eigenvalues against the number of PCs and detect a cut-off beyond which the eigenvalues become constants. Another approach is to discard the PCs with eigenvalues lower than a specified fraction of the first (largest) eigenvalue. There is no simple answer and one has to trade off between errors and the number of PCs for approximation of the observed data which is the primary concern when PCA is used for data compression.

PCA has been applied to analyze functional images including nuclear medicine [72–77] and dynamic MRI [78, 79] where data visualization, structure and functional classification, localization of diseases, and detection of activation patterns are of primary interests. Moeller and Strother [72] applied PCA to analyze functional activation patterns in brain activation experiments. Strother *et al.* [75] revealed an intra- and intersubject subspace in data and demonstrated that the activation pattern is usually contained in the first PC. A later study conducted by Ardekani *et al.* [76] further demonstrated that the activation pattern may spread across several PCs rather than lie only on the first PC, particularly when the number of subjects increases and/or multicenter data are used. PCA was also applied to aid interpretation of oncologic PET images. Pedersen *et al.* [74] applied PCA to aid analyze of dynamic FDG-PET liver data. Anzai *et al.* [77] investigated the use of PCA in detection of tumors in head and neck, also using dynamic FDG-PET imaging. It was found that the first few highest order component images often contained tumors whereas the last several components were simply noise.

Absolute quantification of dynamic PET or SPECT data requires an invasive procedure where a series of blood samples are taken to form an input function for kinetic modeling (Chapter 2 of *Handbook of Biomedical Image Analysis: Segmentation*, *Volume I*). Sampling blood at the radial artery or from an arterialized vein in a hand is the currently recognized method to obtain the input function. However, arterial blood sampling is invasive and has several potential risks associated with both the patient and the personnel who performed the blood sampling [80]. Drawing ROI around vascular structures (e.g., left ventricle in the myocardium [81] and internal carotid artery in the brain [82]) has been

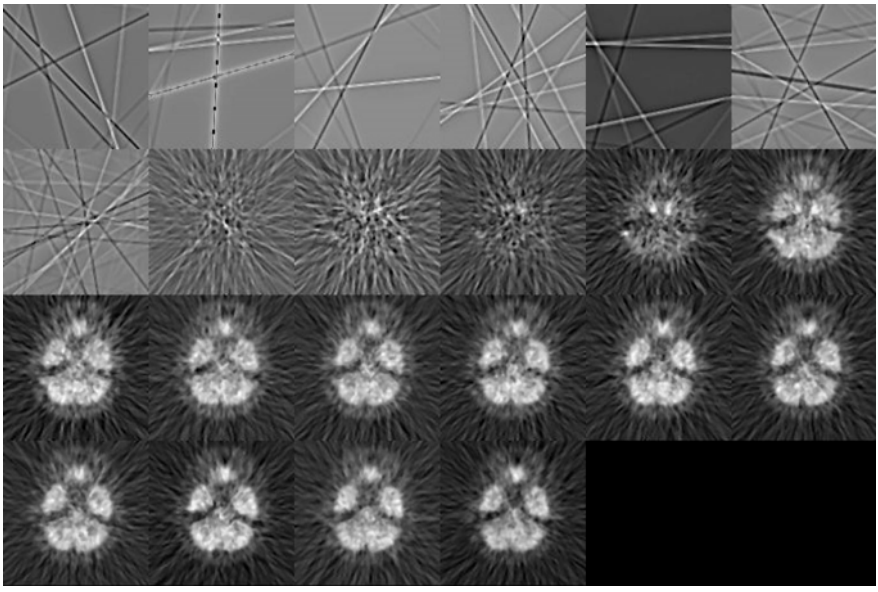


Figure 3.2: A sequence of dynamic neurologic FDG-PET images sampled at the level where the internal carotid arteries are covered. Individual images are scaled to their own maximum.

proposed as a noninvasive method that obviates the need of frequent blood sampling. Delineation of larger vascular structures in the myocardium is relatively straightforward. In contrast, delineation of internal carotid artery in the head and neck is not trivial. A potential application of PCA is the extraction of input function from the dynamic images in which vascular structures are present in the dynamic images. Figure 3.2 show a sequence of dynamic neurologic FDG-PET images sampled at the level in the head where the internal carotid arteries are covered. Figure 3.3 shows the highest 12 PC images. The signal-to-noise ratio (SNR) of the first PC image is very high when comparing with the original image sequence. For PC images beyond the second, they simply represent the remaining variability that the first two PC images cannot account for and they are dominated by noise. The internal carotid arteries can be seen in the second PC image which can be extracted by means of thresholding as mentioned before in sections 3.4.1 and 3.4.4. Figure 3.4 shows a plot of percent contribution to the total variance for individual PCs. As can be seen from the figure, the first and the second PCs contribute about 90% and 2% of the total variance, while the remaining PCs only contribute for less than 0.6% of the total variance individually.

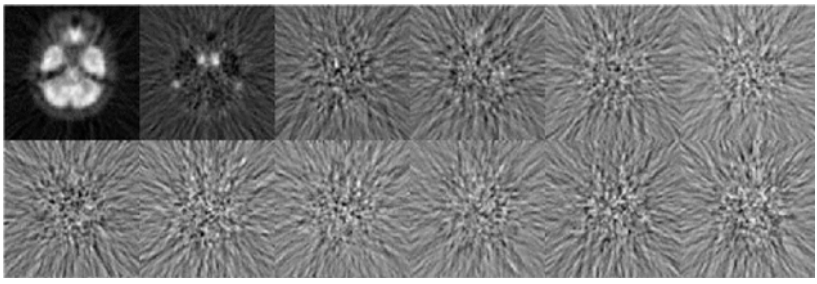


Figure 3.3: From left to right, the figure shows the first six principal component (PC) images (top row), and the 7th to 12th PC images (bottom row) scaled to their own maxima. All but the first two PC images are dominated by noise. The higher order PC images (not shown) look very similar to PC images 3–12.

This means that large amount of information (about 92%) is preserved in only the first two PCs, and the original images can be approximated by making use only the first one or two PCs.

Different from model-led approaches such as compartmental analysis where the physiological parameters in a hypothesized mathematical model are estimated by fitting the model to the data under certain possibly invalid assumptions, PCA is data-driven, implying that it does not rely on a mathematical model.

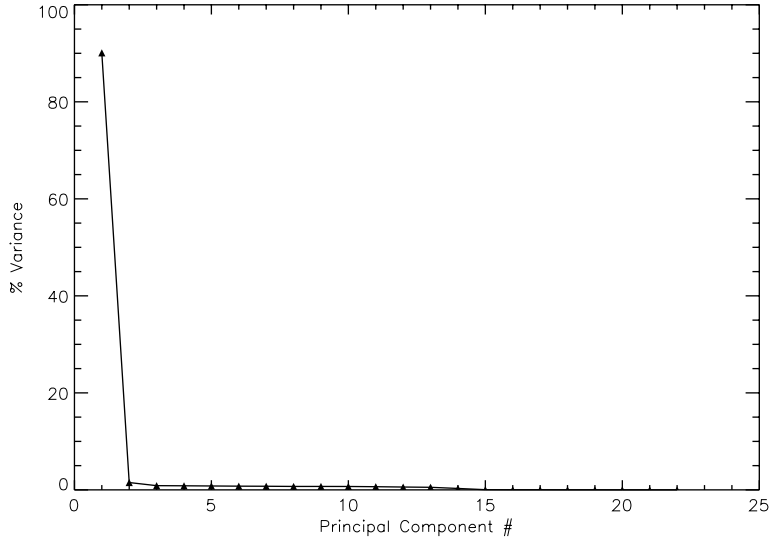


Figure 3.4: The percent variance distribution of the principal component (PC) images.

Instead, it explores the variance–covariance structure of the observed data and finds a set of optimal (uncorrelated) PCs, each of which contains maximal variation present in the measured data. A linear combination of these components can accurately represent the observed data. However, because of lack of model as a constraint, PCA cannot separate signals from statistical noise, which may be an important component if it is highly correlated and dominates the multivariate data. In this case, convincing results of dimensionality reduction or structure exploration may not be achievable as noise is still retained in the higher order components. In addition, the orthogonal components produced by PCA are not necessarily physiological meaningful. Thus, it is difficult to relate the extracted components to the underlying TACs and structures in the multivariate data.

3.5.3.3 Factor Analysis

Factor analysis of dynamic structures (FADS), or *factor analysis* (FA), can be thought of as a generalization of PCA as it produces factors closer to the true underlying tissue response and assumes a statistical model for the observed data. FADS is a semiautomatic technique used for extraction of TACs from a sequence of dynamic images. FADS segments the dynamic sequence of images into a number of structures which can be represented by functions. Each function represents one of the possible underlying physiological kinetics such as blood, tissue, and background in the sequence of dynamic images. Therefore, the whole sequence of images can be represented by a weighted sum of these functions.

Consider a sequence of dynamic images $\mathbf{X}$ of size $M \times N$, with M being the number of pixels in one image and N the number of frames. Each row of $\mathbf{X}$ represents a pixel vector, which is a tissue TAC in PET/SPECT data. Assume that pixel vectors in $\mathbf{X}$ can be represented by a linear combination of factors $\mathbf{F}$, then $\mathbf{X}$ can written as

$$\mathbf{X} = \mathbf{CF} + \boldsymbol{\eta} \tag{3.36}$$

where $\mathbf{C}$ contains factor coefficients for each pixel and it is of size $M \times K$ with K being the number of factors; $\mathbf{F}$ is a $K \times N$ matrix which contains underlying tissue TACs. The additive term $\boldsymbol{\eta}$ in Eq. (3.36) represents measurement noise in $\mathbf{X}$.

Similar to the mathematical analysis detailed before for similarity mapping and PCA, we define $\mathbf{x}_i$ as the ith pixel vector in $\mathbf{X}$, and $\mathbf{f}_k$ being the kth underlying

factor curve (TAC), and c_{ki} being the factor coefficient that represents contribution of the kth factor curve to $\mathbf{x}_i$. Let $\mathbf{Y} = \mathbf{CF}$ and $\mathbf{y}_i$ be a vector which represents the ith row of $\mathbf{Y}$, then

$$\mathbf{y}_i = \sum_{k=1}^{K} c_{ki}\mathbf{f}_k \tag{3.37}$$

and

$$\mathbf{x}_i = \mathbf{y}_i + \boldsymbol{\eta}_i \tag{3.38}$$

where $\boldsymbol{\eta}_i$ represents a vector of noise associated with $\mathbf{x}_i$. Simply speaking, these equations mean that the series of dynamic images $\mathbf{X}$ can be approximated and constructed from some tissue TACs of the underlying structures (represented by a factor model $\mathbf{Y} = \mathbf{CF}$), which are myocardial blood pools, myocardium, liver, and background for cardiac PET/SPECT imaging, for example. The aim of FADS is to project the underlying physiological TACs, $\mathbf{y}_i$ as close as possible to the measured TACs, $\mathbf{x}_i$, so that the MSE between them can be minimized:

$$\Phi(\mathbf{C}, \mathbf{F}) = \sum_{i=1}^{M} \left(\mathbf{x}_i - \sum_{k=1}^{K} c_{ki}\mathbf{f}_k \right)^2 \tag{3.39}$$

Typically, FADS proceeds by first identifying an orthogonal basis for the sequence of dynamic images followed by an oblique rotation. Identification of the orthogonal basis can be accomplished by PCA discussed previously. However, the components identified by PCA are not physiologically meaningful because some components must contain negative values in order to satisfy the orthogonality condition. The purpose of oblique rotation is to impose nonnegativity constraints on the extracted factors (TACs) and the extracted images of factor coefficients [63].

As mentioned in section 3.2, careful ROI selection and delineation are very important for absolute quantification, but manually delineation of ROI is not easy due to high-noise levels present in the dynamic images. Owing to scatter and patient volume effects, the selected ROI may represent "lumped" activities from different adjacent overlapping tissue structures rather than the "pure" temporal behavior of the selected ROI. On the other hand, FADS can separate partially overlapping regions that have different kinetics, and thereby, extraction of TACs corresponding to those overlapping regions is possible.

Although the oblique rotation yield reasonable nonnegative factor curves that are highly correlated with the actual measurements, they are not unique [83] because both factors and factor coefficients are determined simultaneously. It is very easy to see this point by a simple example. Assume that a tissue TAC $\mathbf{x}$ composes of only two factors $\mathbf{f}_1$ and $\mathbf{f}_2$ and c_1 and c_2 being the corresponding factor coefficients. According to Eqs. (3.36) and (3.37), $\mathbf{x}$ can be represented by

$$\mathbf{x} = c_1\mathbf{f}_1 + c_2\mathbf{f}_2 \tag{3.40}$$

which can be written as

$$\mathbf{x} = c_1(\mathbf{f}_1 + \alpha\mathbf{f}_2) + (c_2 - \alpha c_1)\mathbf{f}_2 \tag{3.41}$$

with some constant α. It can be seen that Eqs. (3.40) and (3.41) are equivalent for describing the measured TAC, $\mathbf{x}$, as long as $\mathbf{f}_1 + \alpha\mathbf{f}_2$ and $c_2 - \alpha c_1$ are nonnegative if nonnegativity constraints have to be satisfied. In other words, there is no difference to represent $\mathbf{x}$ using factors $\mathbf{f}_1$ and $\mathbf{f}_2$ and factor coefficients c_1 and c_1, or using factors $\mathbf{f}_1 + \alpha\mathbf{f}_2$ and $\mathbf{f}_2$ and factor coefficients c_1 and $c_2 - \alpha c_1$. Therefore, further constraints such as *a priori* information of the data being analyzed are required [84–87].

FADS has been successfully applied to extract the time course of blood activity in left ventricle from PET images by incorporating additional information about the input function to be extracted [88, 89]. Several attempts have also been made to overcome the problem of nonuniqueness [90, 91]. It was shown that these improved methods produced promising results in a patient planar ^{99m}Tc-MAG$_3$ renal study and dynamic SPECT imaging of ^{99m}Tc-teboroxime in canine models using computer simulations and measurements in experimental studies [90, 91].

3.5.3.4 Cluster Analysis

Cluster analysis has been described briefly in section 3.4.4. One of the major aims of cluster analysis is to partition a large number of objects according to certain criteria into a smaller number of clusters that are mutually exclusive and exhaustive such that the objects within a cluster are similar to each others, while objects in different clusters are dissimilar. Cluster analysis is of potential value in classifying PET data, because the cluster centroids (or centers) are derived

from many objects (tissue TACs) and an improved SNR can be achieved [92]. It has been applied to segment a dynamic [^{11}C]flumazenil PET data [92] and dynamic [^{123}I]iodobenzamide SPECT images [93]. In the following, a clustering algorithm is described. Its application to automatic segmentation of dynamic FDG-PET data for tumor localization and detection is demonstrated in the next section. An illustration showing how to apply the algorithm to generate ROIs automatically for noninvasive extraction of physiological parameters will also be presented.

The segmentation method is based on cluster analysis. Our aim is to classify a number of tissue TACs according to their shape and magnitude into a smaller number of distinct characteristic classes that are mutually exclusive so that the tissue TACs within a cluster are similar to one another but are dissimilar to those drawn from other clusters. The clusters (or clustered ROIs) represent the locations in the images where the tissue TACs have similar kinetics. The kinetic curve associated with a cluster (i.e. cluster centroid) is the average of TACs in the cluster. Suppose that there exists k characteristic curves in the dynamic PET data matrix, $\mathbf{X}$, which has M tissue TACs and N time frames with $k \ll M$ and that any tissue TAC belongs to only one of the k curves. The clustering algorithm then segments the dynamic PET data into k curves automatically based on a weighted least-squares distance measure, $\mathcal{D}$, which is defined as

$$\mathcal{D}\{\mathbf{x}_i, \boldsymbol{\mu}_j\} = \sum_{j=1}^{k} \sum_{i=1}^{M} \|\mathbf{x}_i - \boldsymbol{\mu}_j\|^2_{\mathbf{W}} \tag{3.42}$$

where $\mathbf{x}_i \in \mathbb{R}^N$ is the ith tissue TAC in the data, $\boldsymbol{\mu}_j \in \mathbb{R}^N$ is the centroid of cluster C_j, and $\mathbf{W} \in \mathbb{R}^{N \times N}$ is a square matrix containing the weighting factors on the diagonal and zero for the off-diagonal entries. The weighting factors were used to boost the degree of separation between any TACs that have different uptake patterns but have similar least-squares distances to a given cluster centroid. They were chosen to be proportional to the scanning intervals of the experiment. Although this is not necessarily an optimal weighting, reasonably good clustering results can be achieved.

There is no explicit assumption on the structure of data and the clustering process proceeds automatically in an unsupervised manner. The minimal assumption for the clustering algorithm is that the dynamic PET data can be represented by a finite number of kinetics. As the number of clusters, k, for a given dataset is usually not known *a priori*, k is usually determined by trial and error.

In addition, the initial cluster centroid in each cluster is initialized randomly to ensure that all clusters are nonempty. Each tissue TAC is then allocated to its nearest cluster centroid according to the following criterion:

$$\begin{aligned} & \|\mathbf{x}_l - \boldsymbol{\mu}_i\|^2_{\mathbf{W}} < \|\mathbf{x}_l - \boldsymbol{\mu}_j\|^2_{\mathbf{W}} \\ \Rightarrow \quad & \mathbf{x}_l \in C_i \qquad \forall i, j = 1, 2, \ldots, k, \quad i \neq j \end{aligned} \tag{3.43}$$

where $\mathbf{x}_l \in \mathbb{R}^N$ is the lth tissue TAC in $\mathbf{X}$; $\boldsymbol{\mu}_i \in \mathbb{R}^N$ and $\boldsymbol{\mu}_j \in \mathbb{R}^N$ are the ith and jth cluster centroid, respectively; and C_i represents the ith cluster set. The centroids in the clusters are updated based on Eq. (3.43) so that Eq. (3.42) is minimized. The above allocation and updating processes are repeated for all tissue TACs until there is no reduction in moving a tissue TAC from one cluster to another. On convergence, the cluster centroids are mapped back to the original data space for all voxels. An improved SNR can be achieved because each voxel in the mapped data space is represented by one of the cluster centroids each of which possesses a higher statistical significance than an individual TAC.

Convergence to a global minimum is not always guaranteed because the final solution is not known *a priori* unless certain constraints are imposed on the solution that may not be feasible in practice. In addition, there may be several local minima in the solution space when the number of clusters is large. Restarting the algorithm with different initial cluster centroids is necessary to identify the best possible minimum in the solution space.

The algorithm is similar to the K-means type Euclidean clustering algorithm [40]. However, the K-means type Euclidean clustering algorithm requires that the data are normalized and it does not guarantee that the within-cluster cost is minimized since no testing is performed to check whether there is any cost reduction if an object is moved from one cluster to another.

3.6 Segmentation of Dynamic PET Images

The work presented in this section builds on our earlier research in which we applied the proposed clustering algorithm to tissue classification and segmentation of phantom data and a cohort of dynamic oncologic PET studies [94]. The study was motivated by our on-going work on a noninvasive modeling approach

for quantification of FDG-PET studies where several ROIs of distinct kinetics are required [95, 96]. Manual delineation of ROIs restrain the reproducibility of the proposed modeling technique, and therefore, some other semiautomated and automated methods have been investigated and clustering appears as a promising alternative to automatically segment ROI of distinct kinetics. The results indicated that the kinetic and physiological parameters obtained with cluster analysis are similar to those obtained with manual ROI delineation, as we will see in the later sections.

3.6.1 Experimental Studies

3.6.1.1 Simulated [^{11}C]Thymidine PET Study

To examine the validity of the segmentation scheme, we simulated a dynamic 2-[^{11}C]thymidine (a marker of cell proliferation) PET study. 2-[^{11}C]thymidine was chosen because it is being increasingly used in the research setting to evaluate cancer and treatment response, and it offers theoretical advantages over FDG such as greater specificity in the assessment of malignancy. Also, the kinetics are very similar for most tissues and the data are typically quite noisy. Thus, thymidine data represent a challenging example for testing the clustering algorithm.

Typical 2-[^{11}C]thymidine kinetics for different tissues were derived from eight patients. The data were acquired on an ECAT 931 scanner (CTI/Siemens, Knoxville, TN). The dynamic PET data were acquired over 60 min with a typical sampling schedule (10×30 sec, 5×60 sec, 5×120 sec, 5×180 sec, 5×300 sec) and the tracer TAC in blood was measured with a radial artery catheter following tracer administration. Images were reconstructed using filtered back-projection (FBP) with a Hann filter cut-off at the Nyquist frequency. ROIs were drawn over the PET images to obtain tissue TACs in bone, bone marrow, blood pool, liver, skeletal muscle, spleen, stomach, and tumor. Impulse response functions (IRFs) corresponding to these tissues were determined by spectral analysis of the tissue TACs [97]. The average IRFs for each common tissue type were obtained by averaging the spectral coefficients across the subjects and convolved with a typical arterial input function, resulting in typical TACs for each tissue. The TACs were then assigned to the corresponding tissue types in a single slice of the Zubal phantom [98] which included blood vessels, bone, liver, bone marrow,

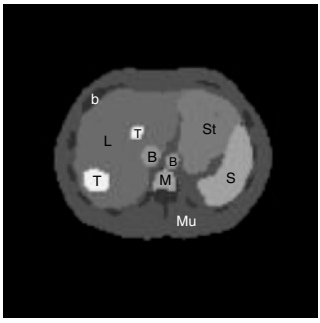


Figure 3.5: A slice of the Zubal phantom. B = blood vessels; b = bone; L = liver; M = marrow; Mu = muscle; S = spleen; St = stomach; T = tumor.

muscle, spleen, stomach, a large and small tumor in the liver (see Fig. 3.5). A dynamic sequence of sinograms was obtained by forward projecting the images into 3.13 mm bins on a 192 $\times$ 256 grid. Attenuation was included in the simulations for the purpose of obtaining the correct scaling of the noise. Poisson noise and blurring were added to simulate realistic sinograms. Noisy dynamic images were then reconstructed using FBP (Hann filter cut-off at the Nyquist frequency). Figure 3.6 shows the metabolite-corrected arterial blood curve and noisy 2-[^{11}C]thymidine kinetics in some representative tissues.

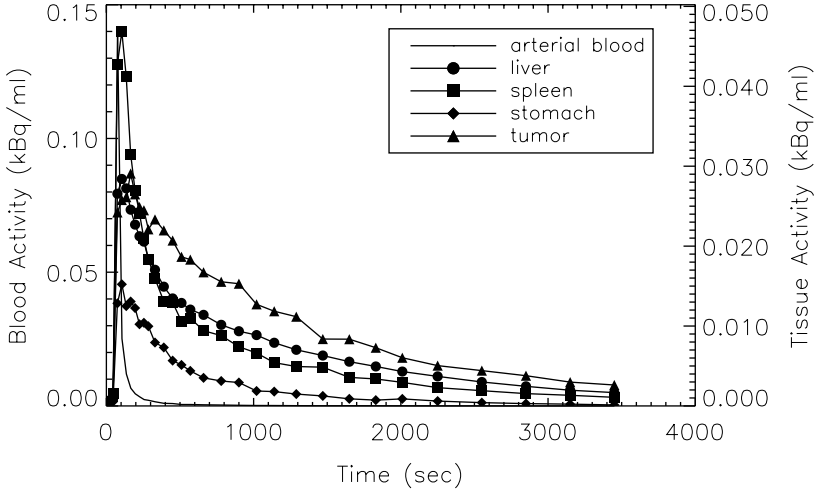


Figure 3.6: Simulated noisy 2-[^{11}C]thymidine kinetics in some representative regions. A metabolite-corrected arterial blood curve, which was used to simulate 2-[^{11}C]thymidine kinetics in different tissues, is also shown.

Figure 3.7: A slice of the Hoffman brain phantom. A tumor in white matter (white spot) and an adjacent hypometabolic region (shaded region) are shown.

3.6.1.2 Simulated FDG-PET Study

Dynamic FDG-PET study was simulated using a slice of the numerical Hoffman brain phantom [99] that modified using a template consisting of five different kinetics (gray matter, white matter, thalamus, tumor in white matter, and an adjacent hypometabolic region in left middle temporal gyrus), as shown in Fig. 3.7. The activities in gray matter and white matter were generated using a five-parameter three-compartment FDG model [100] with a measured arterial input function obtained from a patient (constant infusion of 400 MBq of FDG over 3 min). The kinetics present in the hypometabolic region, thalamus, and tumor were set to 0.7, 1.1, and 2.0 times the activity in gray matter. The kinetics were then assigned to each brain region and a dynamic sequence of sinograms (22 frames, 6×10 sec, 4×30 sec, 1×120 sec, 11×300 sec) was obtained by forward projecting the images into 3.13 mm bins on a 192×256 grid. Poisson noise and blurring were also added to simulate realistic sinograms. Dynamic images were reconstructed using FBP with Hann filter cut-off at the Nyquist frequency. The noisy FDG kinetics are shown in Fig. 3.8 and some of the kinetics are similar to each other due to the added noise and gaussian blurring, although their kinetics are different in the absence of noise and blurring. This is illustrated in the white matter and the hypometabolic region, and the gray matter and thalamus.

3.6.2 Cluster Validation

As mentioned earlier, the optimum number of clusters for a given dataset is usually not known *a priori*. It is advantageous if this number can be determined

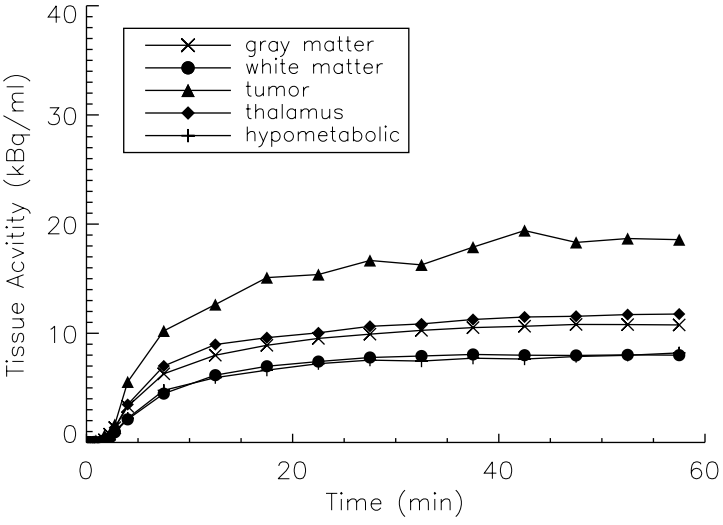


Figure 3.8: Simulated noisy[^{18}F]fluorodeoxyglucose (FDG) kinetics in different regions.

based on the given dataset. In this study, a model-based approach was adopted to cluster validation based on two information-theoretic criteria, namely, Akaike information criterion (AIC) [101] and Schwarz criterion (SC) [102], assuming that the data can be modeled by an appropriate probability distribution function (e.g. Gaussian). Both criteria determine the optimal model order by penalizing the use of a model that has a greater number of clusters. Thus, the number of clusters that yields the lowest value for AIC and/or SC is selected as the optimum. The use of AIC and SC has some advantages compared to other heuristic approaches such as the "bootstrap" resampling technique which requires a large amount of stochastic computation. This model-based approach is relatively flexible in evaluating the goodness-of-fit and a change in the probability model of the data does not require any change in the formulation except the modeling assumptions. It is noted, however, that both criteria may not indicate the same model as the optimum [102].

The validity of clusters is also assessed visually and by thresholding the average mean squared error (MSE) across clusters, which is defined as

$$\text{MSE} = \frac{1}{k} \sum_{j=1}^{k} \sum_{i=1}^{M} \|\mathbf{x}_i - \boldsymbol{\mu}_j\|_{\mathbf{W}}^2. \tag{3.44}$$

Both approaches are subjective but they can provide an insight into the "correct" number of clusters.

3.6.3 Human Studies

The clustering algorithm has been applied to a range of FDG-PET studies and three examples (two patients with brain tumor and one patient with a lung cancer) are presented in this chapter. FDG-PET was chosen to assess the clustering algorithm because it is commonly used in clinical oncologic PET studies. All oncological PET studies were performed at the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, Australia. Ethical permissions were obtained from the Institutional Review Board.

Dynamic neurologic FDG-PET studies were performed on an ECAT 951R whole-body PET tomograph (CTI/Siemens, Knoxville, TN). Throughout the study the patient's eyes were patched and ears were plugged. The patients received 400 MBq of FDG, infused at a constant rate over a 3-min period using an automated injection pump. At least 30 min prior to the study, patient's hands and forearms were placed into hot water baths preheated to 44 °C to promote arterio-venous shunting. Blood samples were taken at approximately 30 sec for the first 6 min, and at approximately 8, 10, 15, 30, and 40 min, and at the end of emission data acquisition. A dynamic sequence of 22 frames was acquired for 60 min following radiotracer administration according to the following schedule: 6×10 sec, 4×30 sec, 1×2 min, 11×5 min. Data were attenuation corrected with a postinjection transmission method [103]. Images were reconstructed on a 128×128 matrix using FBP with a Shepp and Logan filter cut-off at 0.5 of the Nyquist frequency.

The dynamic lung FDG-PET study was commenced after intravenous injection of 487 MBq of FDG. Emission data were acquired on an ECAT 951R whole-body PET tomograph (CTI/Siemens, Knoxville, TN) over 60 min (22 frames, 6×10 sec, 4×30 sec, 1×2 min, and 11×5 min). Twenty one arterial blood samples were taken from the pulmonary artery using a Grandjean catheter to provide an input function for kinetic modeling.

The patient details are as follows:

Patient 1: The FDG-PET scan was done in a female patient, 6 months after resection of a malignant primary brain tumor in the right parieto-occipital

lobe. The scan was done to determine if there was evidence for tumor recurrence. A partly necrotic hypermetabolic lesion was found in the right parieto-occipital lobe that was consistent with tumor recurrence.

Patient 2: A 40-year-old woman had a glioma in the right mesial temporal lobe. The FDG-PET scan was performed at 6 months after tumor resection. A large hypermetabolic lesion was identified in the right mesial temporal lobe that was consistent with tumor recurrence.

Patient 3: A 67-year-old man had an aggressive mesothelioma in the left lung. In the PET images, separate foci of increased FDG uptake were seen in the contralateral lymph nodes as well as in the peripheral left lung.

As they are unnecessary for clustering and the subsequent analysis, low count areas such as the background (where the voxel values should be zero theoretically) and streaks (which are due to reconstruction errors) were excluded by zeroing voxels whose summed activity was below 5% of the mean pixel intensity of the integrated dynamic images. A 3×3 closing followed by a 3×3 erosion operation was then applied to fill any "gap" inside the intracranial/body region to which cluster analysis was applied. Parametric images of the physiological parameter, K, which is defined as the value of $k_1^* k_3^* / (k_2^* + k_3^*)$ [104], were generated by fitting all voxels inside the intracranial/body region using Patlak graphical approach [105]. The resultant parametric images obtained for the raw dynamic images and dynamic images after cluster analysis were assessed visually. Compartmental model fitting using the three-compartment FDG model [104] was also performed on the tissue TACs extracted manually and by cluster analysis to investigate whether there is any disagreement between the parameter estimates.

3.6.4 Results

3.6.4.1 Simulated [^{11}C]Thymidine PET Study

Figure 3.9 shows the segmentation results using different numbers of clusters, k, in the clustering algorithm. The number of clusters is actually varied from 3 to 13 but only some representative samples are shown. In each of the images in Figs. 3.9(a)–3.9(f), different gray levels are used to represent the cluster locations. Figure 3.9 shows that when the number of clusters is small, segmentation

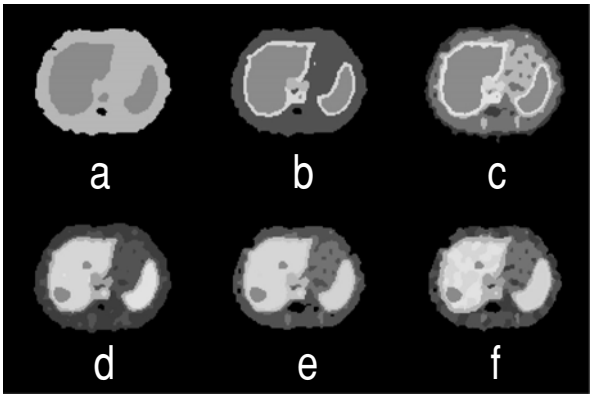


Figure 3.9: Tissue segmentation obtained with different number of clusters. (a) $k = 3$, (b) $k = 5$, (c) $k = 7$, (d) $k = 8$, (e) $k = 9$, and (f) $k = 13$. (Color Slide)

of the data is poor. With $k = 3$, the liver, marrow, and spleen merge to form a cluster and the other regions merge to form a single cluster. With $5 \leq k \leq 7$, the segmentation results improve because the blood vessels and stomach are visualized. However, the hepatic tumors are not seen and the liver and spleen are classified into the same cluster. With $k = 8$, the tumors are visualized and almost all of the regions are correctly identified (Fig. 3.9(d)). Increasing the value of k to 9 gives nearly the same segmentation as in the case of $k = 8$ (Fig. 3.9(e)). Further increasing the value of k, however, may result in poor segmentation because the actual number of tissues present in the data is less than the specified number of clusters. Homogeneous regions are therefore fragmented to satisfy the constraint on the number of clusters (Fig. 3.9(f)). Thus, 8 or 9 clusters appear to provide reasonable segmentation of tissues in the slice and this number agrees with the various kinetics present in the data.

Figure 3.10 plots the average MSE across clusters as a function of k. The average MSE decreases monotonically, as it drops rapidly ($k < 8$) before reaching a plateau ($k \geq 10$). From the trend of the plot, there is no significant reduction in the average MSE with $k > 12$. Furthermore, the decrease in the average MSE is nearly saturated with $k \geq 8$. These results confirm the findings of the images in Fig. 3.9, suggesting 8 or 9 as the optimal number of clusters for this dataset.

Table 3.1 tabulates the results of applying AIC and SC to determine the optimum number of clusters which is the one that gives the minimum value for

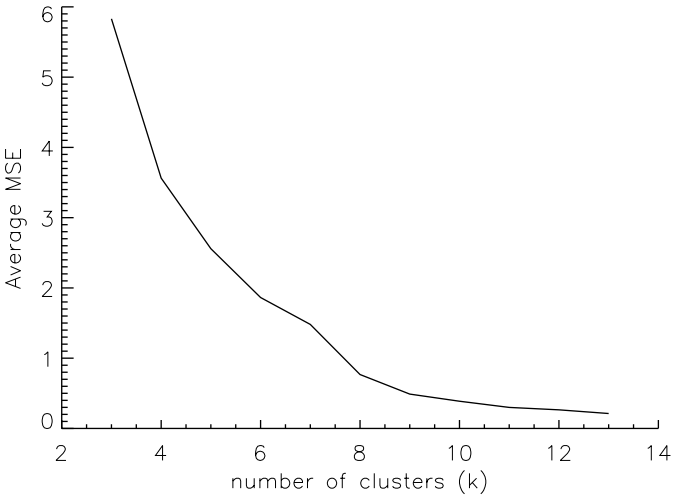


Figure 3.10: Average mean squared error (MSE) as a function of number of clusters.

the criteria. Both criteria indicate that $k = 8$ is an optimal approximation to the underlying number of kinetics. It was found that a good segmentation can be achieved when the number of clusters is the same as that determined by the criteria. Conversely, the segmentation result is poor when the number of clusters is smaller than that suggested by the criteria and there is no significant improvement in segmentation when the number of clusters is larger than that determined by the criteria. The heuristic information given by both criteria also support our visual interpretation of the clustering results, suggesting that the criteria are reasonable approaches to objectively determine the number of clusters.

Table 3.1: Computed values for AIC and SC with different choices of the value of k

	Number of clusters, k										
Criterion	3	4	5	6	7	8	9	10	11	12	13
AIC	99005	95354	93469	90904	88851	**86967**[a]	89769	93038	91994	90840	89807
SC	98654	94888	92887	90206	88038	**86038**[a]	88725	91878	90719	89450	88301

AIC: Akaike information criterion; SC: Schwarz criterion.

[a] Values in bold correspond to the computed minimum of the criterion.

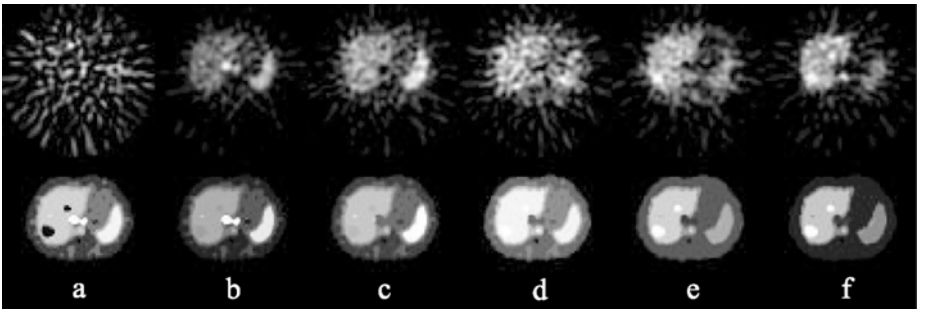


Figure 3.11: Single slice of simulated 2-[^{11}C]thymidine PET study. Top row shows the original reconstructed images at (a) 15 sec, (b) 75 sec, (c) 135 sec, (d) 285 sec, (e) 1020 sec, and (f) 2850 sec postinjection. Bottom row shows same slice at identical time points after cluster analysis. Individual images are scaled to their own maximum.

Application of the clustering algorithm to the simulated PET data is shown in Fig. 3.11. The number of clusters is eight, corresponding to the optimum number of clusters determined by the statistical criteria. The SNR of the images is markedly improved after clustering. In addition, the blood vessels are clearly seen in the frame sampled at 15 and 75 sec after clustering but not in the corresponding frame in the original data. In the original images, it is difficult to identify different tissues which may be due to reconstruction effects and inhomogeneous noise. However, the liver, spleen, muscle, marrow, stomach, and tumors are clearly delineated by the clustering algorithm (bottom row of the figure).

3.6.4.2 Simulated FDG-PET Study

Five cluster images were generated by applying the clustering algorithm to the noisy simulated dynamic images. The number of clusters k was actually varied from 3 to 10 and the optimal k was determined by inspecting the change of average MSE and the visual quality of the cluster images. Figure 3.12 shows the cluster images for $k = 5$ that was found to be the optimum number of clusters for this dataset. It was found that the tumor cannot be located when k was small ($k < 4$). The tumor was located by gradually increasing the number of clusters. However, there was a deteriorated segmentation of all regions when k was large ($k > 7$).

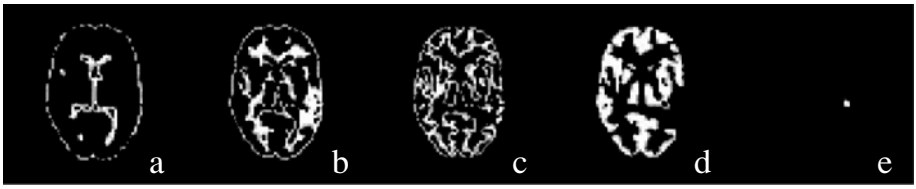


Figure 3.12: Five cluster images obtained from the noisy simulated dynamic FDG-PET data. The images correspond to (a) ventricles and scalp, (b) white matter and left middle temporal gyrus (hypometabolic zone), (c) partial volume between gray matter and white matter, (d) gray matter, deep nuclei, and outer rim of tumor, and (e) tumor.

Although the tumor was small in size, cluster analysis was still able to locate it because of its abnormal temporal characteristics as compared to the other regions. Cluster analysis also performed well in extracting underlying tissue kinetics in gray matter and white matter because of their distinct kinetics. On the contrary, the kinetics in the thalamus and the hypometabolic region were not separated from those in gray and white matter but this was not unexpected since their kinetics were very similar.

Owing to the partial volume effects (PVEs), there were some vague regions whose kinetics were indeterminate (Fig. 3.12(c)) and did not approach gray or white matter. The algorithm was unlikely to assign such kinetics to the cluster corresponding to white matter or to the cluster corresponding to gray matter since the overall segmentation results would be deteriorated. Thus, a cluster was formed to account for the indeterminate kinetics.

3.6.4.3 Human Studies

Segmentation results are shown for dynamic neurologic (Fig. 3.13) and lung (Fig. 3.14) FDG-PET studies. The clusters are represented by differing gray scales and slices were sampled at the level where the lesions were seen on the original reconstructed data. Since there is no *a priori* knowledge about the optimum number of clusters, the value of k was varied in order to determine the optimal segmentation using the AIC and SC as in the phantom study. For Fig. 3.13, eight clusters were found to give the optimal segmentation for these datasets. The locations of the tumors and the rim of increased glucose uptake

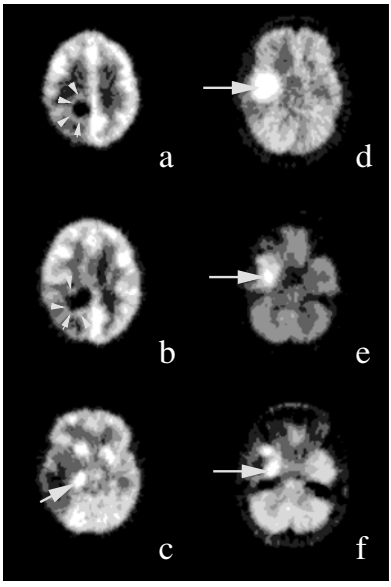


Figure 3.13: Tissue segmentation obtained from Patient 1 at (a) slice 10, (b) slice 13, and (c) slice 21; and Patient 2 at (d) slice 21, (e) slice 24, and (f) slice 26. The number of clusters used is eight. The locations of the solid hypermetabolic portions of the tumors (arrows) and the small rim of increased glucose uptake (arrow heads) identified by cluster analysis are shown.

are identified correctly by the clustering algorithm with the optimal value of clusters.

For Fig. 3.14, the number of clusters was varied from 3 to 13 and only some representative results are shown. Similar to the simulation study, the segmentation results are poor when the number of clusters is small ($k = 3$), while the segmentation is gradually improved by increasing the number of clusters. Based on the AIC and SC, the optimum numbers of clusters for the selected slices (4, 19, and 24) were found to be 8, 8, and 9, respectively. It is not surprising that the optimum number of clusters is different for different slices because of the differing number of anatomical structures contained in the plane and the heterogeneity of tracer uptake in tissues. Nevertheless, the tumor (slice 4), right lung and muscle (slices 4, 19, and 24), blood pool (slices 4, 19, and 24), separate foci of increased FDG uptake (slices 19 and 24), and the injection site (slices 4, 19, and 24) are identifiable with the optimum number of clusters.

Figure 3.15 shows the measured blood TAC at the pulmonary artery and the extracted tissue TACs for the tumor (from slice 4), lung and muscle (from

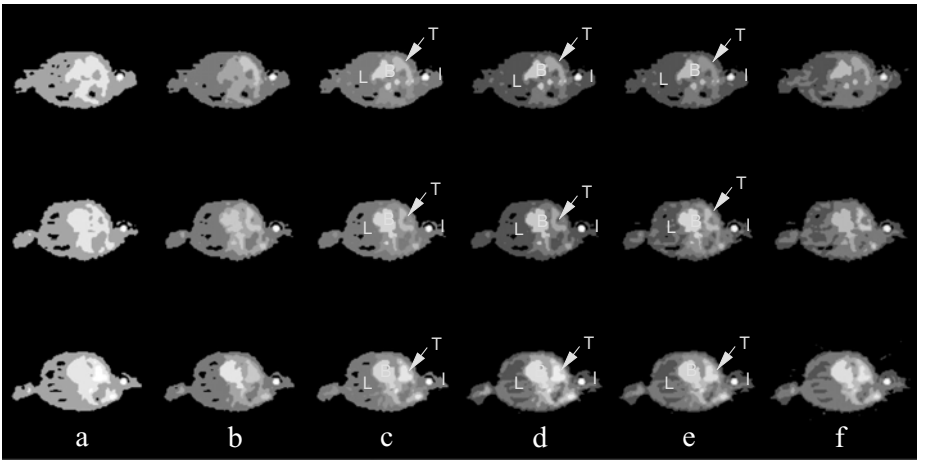


Figure 3.14: Tissue segmentation of the dynamic lung FDG-PET data from Patient 3 in three selected slices: 4 (top row), 19 (middle row), and 24 (bottom row) with different number of clusters. (a) $k = 4$, (b) $k = 7$, (c) $k = 8$, (d) $k = 9$, (e) $k = 10$, and (f) $k = 12$. (I = injection site; B = blood pool; L = lung; T = tumor).

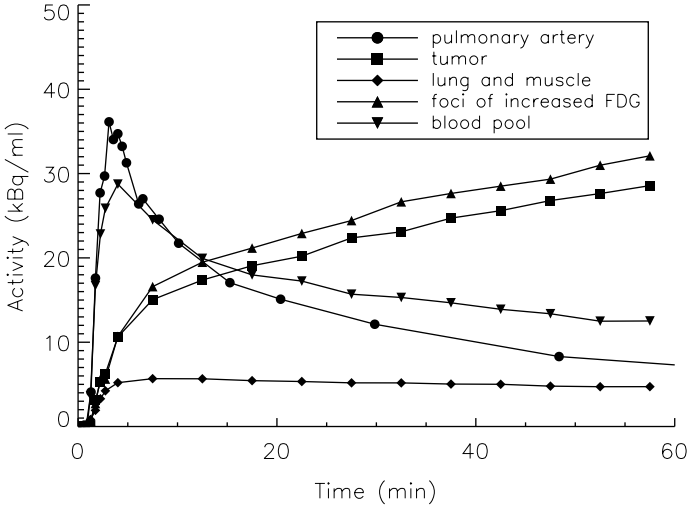


Figure 3.15: Extracted tissue time–activity curve (TACs) corresponding to the tumor, lung, and muscle, foci of increased FDG uptake, and blood pool. The measured blood TAC at the pulmonary artery is also shown.

Table 3.2: Compartmental modeling of the tumor TACs obtained by manually ROI delineation and by cluster analysis

Parameter	Manual delineation	Cluster analysis
k_1^* (ml/min/g)	0.854 ± 17.1	0.921 ± 18.2
k_2^* (min^{-1})	1.987 ± 21.2	2.096 ± 22.5
k_3^* (min^{-1})	0.099 ± 0.9	0.100 ± 0.9
k_4^* (min^{-1})	0.018 ± 1.1	0.017 ± 1.2
K (ml/min/g)	0.041 ± 5.3	0.042 ± 5.3

TAC: Time–activity curve; ROI: region of interest. Values are given as estimate ±% CV.

slice 19), foci of increased FDG uptake (from slice 24), and the blood pool (from slice 19) using the corresponding optimal value of clusters.

The extracted tissue TACs obtained by cluster analysis and manual ROI delineation were fitted to the three-compartment FDG model using nonlinear least squares method and the results obtained for the tumor tissue TAC (Patient 2) are summarized in Table 3.2. There was a close agreement between the parameter estimates for the tissue TACs obtained by different methods in terms of the estimate and the coefficient of variation (CV), which is defined as the ratio of the standard deviation of the parameter estimate to the value of the estimate. Similar results were also found for other regions.

3.7 Extraction of Physiological Parameters and Input Function

Quantification of dynamic PET or SPECT data requires an invasive procedure where a series of blood samples are taken to form an input function for kinetic modeling. One of the potential applications of the clustering algorithm presented earlier is in noninvasive quantitative PET. We have proposed a simultaneous estimation approach to estimate the input function and physiological parameters simultaneously with two or more ROIs and our results with *in vivo* PET data are promising [95]. The method is still limited, however, by the selection of ROIs whose TACs must have distinct kinetics. As the ROIs are drawn manually on the PET images, reproducibility is difficult to achieve. The use of clustering to extract tissue TACs of distinct kinetics has been investigated in three

Table 3.3: Comparison between the estimated input functions obtained using different number of *manually drawn* ROIs and *clustered* ROIs, and the measured input functions

	Number of ROIs			
	2	3	4	5
Manually drawn ROIs				
MSE	0.632	0.365	0.431	0.967
AUC (measured = 24.077)	23.796	23.657	24.188	25.138
Linear regression on AUC ($n = 19$)				
Slope	0.967	0.963	0.984	1.022
Intercept	0.493	0.460	0.609	0.712
r value	0.999	0.999	0.999	0.999
Clustered ROIs				
MSE	0.100	0.096	0.040	0.066
AUC (measured = 24.077)	20.742	23.721	25.430	23.481
Linear regression on AUC ($n = 19$)				
Slope	0.807	0.953	1.067	0.946
Intercept	0.874	0.575	-0.321	0.481
r value	0.993	0.998	0.999	0.999

MSE = Mean square errors between the estimated and the measured input functions; AUC = area under the blood curve; r = coefficient of correlation; ROI = region of interest.

FDG-PET studies. Table 3.3 summarizes the results for the estimation of the input functions by the proposed modeling approach for both manually drawn ROIs and clustered ROIs. The MSE between the estimated and the measured input functions are tabulated. In addition, results of linear regression analysis on the areas under the curves (AUCs) covered by the measured and the estimated input functions are listed for comparison. Regression lines with slopes close to unity and intercepts close to zero were obtained in all cases for manually drawn ROIs and for clustered ROIs.

Figure 3.16 plots the measured input function and the estimated input functions for manually drawn ROIs and clustered ROIs, respectively. The estimated input functions were obtained by simultaneously fitting with three ROIs of distinct kinetics. There was a very good agreement between the estimated input functions and the measured blood curve, in terms of the shape and the peak time estimation at which the peak occurs, despite the overestimation of the peak value. Thus, cluster analysis may be useful as a preprocessing step before our noninvasive modeling technique.

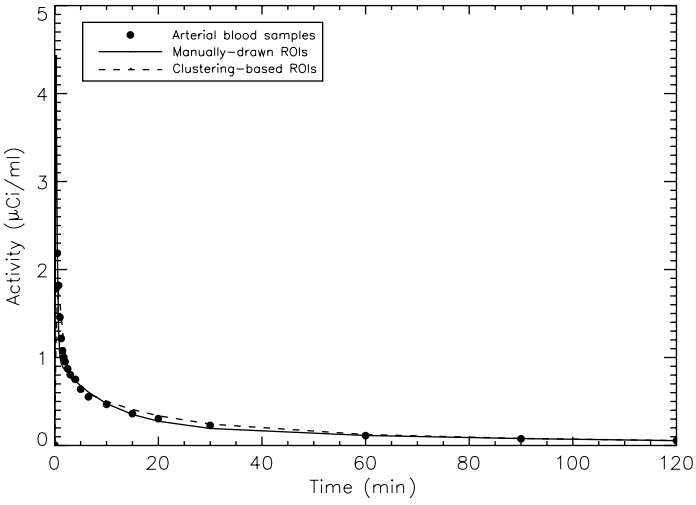


Figure 3.16: Plot of the measured arterial input function, the estimated input function from *manually drawn regions of interest (ROIs)*, and *clustering based ROIs*. The estimated input functions were obtained by simultaneously fitting with three ROIs of distinct kinetics.

Alternatively, clustering can be applied to extract input function directly on the dynamic PET/SPECT data if the vascular structures (e.g. left ventricle [81] and internal carotid artery [82]) are present in the field of view, providing that partial volume and spillover effects are appropriately corrected. Clustering has also been found useful in analyzing PET/SPECT neuroreceptor kinetics in conjunction with simplified techniques for quantification [106]. In particular, identification of regions that are devoid of specific binding is attractive because the kinetics of these regions can be treated as a noninvasive input function to the simplified approach for parametric imaging of binding potentials and relative delivery [107, 108].

3.8 Fast Generation of Parametric Images

Fast generation of parametric images is now possible with current high-speed computer workstations. However, overestimation of parameters and negative parameter estimates, which are not physiologically feasible, occur often when the data are too noisy. Reliable parametric imaging is therefore largely dependent

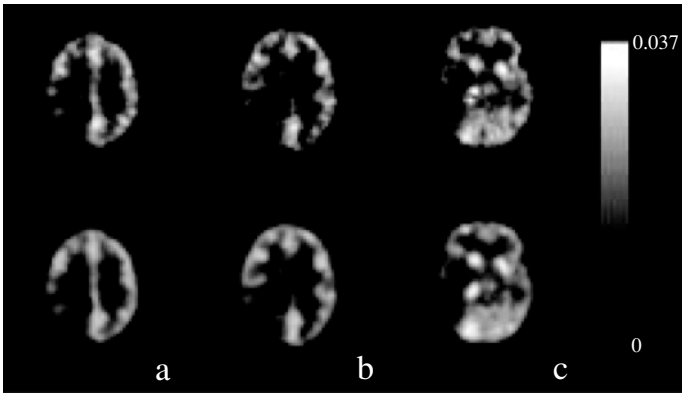


Figure 3.17: Parametric images on a pixel-by-pixel basis of K obtained from Patient 1: (a) slice 10; (b) slice 13; (c) slice 21. Top row shows the images obtained from the raw dynamic images and bottom row shows the images obtained from dynamic images after cluster analysis. The images have been smoothed slightly for better visualization.

on the noise levels inherent in the data which affect, in addition to meaningful parameter estimation, the time required to converge as well as the convergence. Clustering may be useful as a preprocessing step before fast generation of parametric images since only a few characteristic curves which have high statistical significance, need to be fitted as compared to conventional pixel-by-pixel parametric image generation where many thousands of very noisy tissue TACs must be analyzed. The computational advantage and time savings for generation of parametric images (fitting many thousands of kinetic curves versus several curves) are apparent.

Figure 3.17 shows the parametric images of physiological parameters, K, obtained from the neurologic study for Patient 1 in the three selected slices. The top and bottom rows of the images correspond to the results obtained from pixel-by-pixel fitting the TACs in the raw dynamic PET data and data after cluster analysis, respectively. The K images are relatively noisy when compared to the data after cluster analysis because of the high-noise levels of PET data which hampered reliable parametric image generation. However, the visual quality of the K images improves markedly with cluster analysis as a result of the increased SNR of the dynamic images. Low-pass filtering of the original parametric images may improve the SNR but clustering should produce better results because it takes

the tissue TACs with similar temporal characteristics for averaging. Meanwhile, low-pass filtering only makes use of the spatial (adjacent pixels) information for filtering and this will only further degrade the spatial resolution. The feasibility of using the kinetic curves extracted by cluster analysis for noninvasive quantification of physiological parameters and parametric imaging has been investigated and some preliminary data have been reported [109]. Some other recent studies can be found elsewhere [110–115].

3.9 Visualization of Disease Progress

In nuclear medicine, several kinds of organ function can be measured simultaneously with various radiopharmaceuticals under different conditions. This gives us useful information about the stage of disease progress if the relationship between various parameters such as metabolism, blood flow, and hemodynamics can be elucidated. Toyama *et al.* [116] investigated the use of agglomerative hierarchical and K-means clustering methods to study regional vasodilative and vascoconstrictive reactivity and oxygen metabolism before and after revascularization surgery in chronic occlusive cerebrovascular disease. By clustering a four-variable correlation map, whose pixel values on the X, Y, Z, and T axes represent, respectively, the resting cerebral blood flow, the hyperventilatory response, the acetazolamide response, and regional oxygen metabolic rate, anatomically and pathophysiologically different areas can be identified showing the involvement of certain areas with varying degrees of progression between pre- and postsurgical treatment, while functional changes in the revascularized region can be depicted on the clustered brain images. It appears that clustering technique maybe useful for multivariate staging of hemodynamic deficiency in obstructive cerebrovascular disease and it also be suitable for objective representation of multiple PET physiological parameters obtained from ^{15}O-labeled compound studies and also in brain activation studies.

3.10 Characterization of Tissue Kinetics

Kinetic modeling of radiotracer (or radiopharmaceutical) is the core of dynamic PET/SPECT imaging. The aim of modeling is to interpret kinetic data

quantitatively in terms of physiological and pharmacological parameters of a mathematical model, which describes the exchanges (e.g. delivery and uptake) of radiotracer by the tissue. Statistical inferences can then be made regarding the distribution and circulation of tracers within different tissues regions, which are quantitatively represented by the physiological/pharmacological parameters in the model. Successful statistical inference relies heavily on the appropriate use of analysis approaches and *a priori* knowledge of the underlying system as well as the validity of the assumptions being made. What if we know nothing about the underlying system, or little is known about the tracer characteristics and we are unsure if the assumptions (e.g. tissue homogeneity) being made are valid? The use of kinetic modeling could lead to incorrect inferences about the complex physiological or biochemical processes. In this case, data-driven approaches can provide important clues to what is going on inside the underlying system and how the radiotracer behaves inside the tissue, as they interrogate the measured data to characterize the complex processes, with minimal assumptions and independent of any kinetic model.

Evaluation of soft tissue sarcomas (STS) is a challenging clinical problem because these tumors are very heterogeneous, and the treatment of patients with STS is also very complicated. The most essential step in the diagnostic evaluation of STS is tumor biopsy. PET imaging has the ability to differentiate benign from malignant lesions. It can detect intralesional morphologic variation in soft tissue sarcomas, and it is of value in grading tumor, staging, restaging, and prognosis. Fluoromisonidazole (FMISO) has been shown to bind selectively to hypoxic cells *in vitro* and *in vivo* at radiobiologically significant oxygen levels. When labeled with the positron emitter fluorine-18 (^{18}F), its uptake in tissue can be localized and detected quantitatively with high precision by PET. [^{18}F]FMISO uptake has been investigated in various human malignancies [117]. PET imaging with [^{18}F]FMISO, FDG, and ^{15}O-water may provide valuable information complementary to tumor biopsy for better understanding of the biological behavior of STS. As cluster analysis does not rely on tracer assumptions nor kinetic model, it maybe useful in analyzing tissue TACs of STS obtained from [^{18}F]FMISO-PET and FDG-PET, and in looking for any correlations, for instance, tumor volume, hypoxic volume, and vascular endothelial growth factor expression, etc, between different datasets [118].

3.11 Partial Volume Correction in PET

Ideally, after corrections for the physical artifacts (e.g. scatter and attenuation) and calibration, the reconstructed PET image should represent highly accurate radiopharmaceutical distribution in absolute units of radioactivity concentration throughout the field of view of the PET scanner. However, this only holds for organs or structures with dimensions greater than twice the spatial resolution of the scanner, which is characterized by the full width at half maximum height of an image of a line source. When the object or structure being imaged is smaller than this, the apparent activity concentration is diluted. The degree of dilution in activity concentration varies with the size of the structure being imaged and the radioactivity concentration of the imaged structure comparing to its surrounding structures [10]. This phenomenon is known as *partial volume effect* (PVE), which is solely caused by the limited spatial resolution of the PET scanner.

A number of approaches have been proposed to correct or minimize the PVE, including resolution recovery before or during image reconstruction, and incorporation of side information provided from anatomical imaging modalities such as CT and MRI. One of the popular approaches that incorporates MRI segmentation for partial volume correction is the method proposed by Müller-Gärtner *et al.* [54] but the method is only applicable to brain imaging. PET images are first spatially co-registered with MR images obtained from the same subject. The MR images are then segmented into gray matter, white matter, and CSF regions, represented in three separate images. These images are then convolved spatially with a smoothing kernel which is derived from the point spread function of the PET scanner. The convolved white matter, image is then normalized to the counts in a white matter ROI drawn on the PET image so that spillover white matter activity into gray matter regions can be removed. Finally, the resultant image is divided by the smoothed gray matter image so that signals in small structures, which were smoothed severely, can be enhanced.

3.12 Attenuation Correction in PET

Accurate attenuation correction (AC) is essential to emission computed tomography such as PET and SPECT, for both quantitative and qualitative

interpretation of the results (Chapter 2 of *Handbook of Biomedical Image Analysis: Segmentation*, *Volume I*). In PET, for instance, AC factors are most often determined by calculating the pixel-wise ratio of a blank scan acquired before positioning the patient in the gantry of the scanner, and a transmission scan performed with the patient in the gantry. The major drawback of this approach is that statistical noise in the transmission data would propagate to the emission data [119, 120]. Depending on several factors such as body size and composition, transmission scans of 15–30 min are often performed to minimize the propagation of noise to the emission data through AC, at the price of reducing the patient throughput and increasing the errors due to patient motion, causing misalignment between transmission and emission data. Segmented AC methods, which employ image segmentation to partition the transmission images into regions of homogeneous density such as soft tissue, lung, and solid bones whose AC factors are known *a priori* and can be assigned, are particularly useful in cases where propagation of noise in transmission measurements during AC becomes a significant effect. A number of approaches based on the framework of pixel classification techniques and region-based segmentation approaches have been proposed and examined for segmented AC in PET. For example, Huang *et al.* [121] proposed a method where the operator manually defines the outlines of the body and the lung on the attenuation images. Known attenuation coefficients are then assigned to these regions and noiseless AC factors are then obtained by forward projecting the attenuation images with assigned attenuation coefficients. This approach has been further extended by a number of investigators by automating the determination of lung and the body regions using image segmentation techniques. For instance, Xu *et al.* used local thresholding approach to segment attenuation images into air, lung, and soft tissue [122]. Meikle *et al.* [123] used histogram fitting techniques to assign the attenuation values based on an assumed probability distribution for the lung and soft tissue components. Papenfuss *et al.* [124] used expectation-maximization clustering technique in conjunction with thresholding to produce fuzzy segmentation of attenuation images. Likewise, Bettinardi *et al.* [125] proposed an adaptive segmentation technique, also based on fuzzy clustering. This method can automatically determine the number of tissue classes in the attenuation images. The method can generally be applied to any region of the body. At least one of the aforementioned methods is currently in routine use by many PET centers worldwide.

3.13 Application to Analysis of fMRI Data

Functional MRI is a powerful modality for determining neural correlates of cognitive processes. It can be used to monitor changes of physiological parameters such as regional cerebral blood flow, regional cerebral blood volume, and blood oxygenation during cognitive tasks [126]. To extract functional information and detect activated regions using fMRI, the most widely adopted procedures are generally based on statistics theory and are paradigm dependent [65, 127, 128]. Cluster analysis has recently been applied to the fMRI discipline [129–131]. It is anticipated that cluster analysis will have great impact on analysis of fMRI signals for the detection of functional activation sites during cognitive tasks.

3.14 Discussion and Conclusions

In this chapter, a number of segmentation techniques used in, but not specific to functional imaging have been detailed. In particular, tissue segmentation and classification in functional imaging are of primary interests for dynamic imaging, for which cluster analysis is a valuable asset for data analysis because the identified characteristic curves are in the same space as the original data. This certainly has advantages over PCA in terms of interpretation of identified PCs, and over FADS where the factor components are rotated, leading to possibly nonunique factor explanation and interpretation. This chapter focuses on functional segmentation and a clustering technique is presented and discussed in great detail. The proposed technique is an attempt to overcome some of the limitations associated with commonly used manual ROI delineation, which is labor intensive and time-consuming. The clustering technique described is able to provide statistically meaningful clusters because the entire sequence of images are analyzed and different kinetic behaviors and the associated regions are extracted from the dataset, as long as there is a finite number of kinetics in the data. Once the segmentation process is completed, the extracted TACs, i.e. the cluster centroids, are then mapped back to the original data space for all voxels. Thus, an improved SNR can be achieved because each voxel in the mapped data space is represented by one of the cluster centroids each of which possesses a higher statistical significance than an individual TAC in the same spatial

location. Therefore, the extracted TACs obtained by cluster analysis should be more consistent and reproducible.

It is difficult to identify obvious cluster centroids in PET data because they are multidimensional and noisy. Therefore, initial centroids are needed for the proposed algorithm. The initial cluster centroids do not have to be accurate because they are used only as seeds to start the algorithm. However, if the starting centroids are far from the final cluster centroids, more iterations may be required. An incorrect initial selection may occur if a noisy outlier is chosen, resulting in a cluster with a single member. For this case, a lower bound on the final number of members in a cluster should be incorporated to prevent the cluster from being exhausted.

The optimum number of clusters for cluster analysis is usually not known *a priori*. The number of clusters, k, is also dependent on a number of factors mentioned previously. In addition, different choices for the values of k may result in different partitions of data. In this study, we limited the range for the values of k and applied the clustering algorithm to the simulated and real data. Nevertheless, it is reasonable to assume that the limited number of clusters used in this study is feasible, given that there is a finite number of kinetics present in the data. With the use of information-theoretic approaches to cluster validation, one can objectively determine the optimum number of clusters for the given dataset. However, caution should be taken when using the criteria as they are model dependent. The optimum number of clusters suggested by the criteria may not make sense if the specified probability distribution function for the observed data is not appropriate. There are a number of statistical criteria for the determination of the optimal number of clusters in addition to those used in this study and we are currently exploring various approaches, including the minimum description length to the cluster validation problem.

A limitation of the proposed algorithm is that it cannot differentiate anatomical structures having similar kinetics but are unconnected spatially. It is expected that future work will consider the addition of other information such as the geometry and the coordinates of the structure concerned. Another related issue is tissue heterogeneity [132] although this effect is usually ignored. In the work described, we did not attempt to solve this problem for cluster analysis. However, some heuristic interpretations could be made. In anatomy, most of the anatomical structures are discrete and well separated, they should easily be segmented by the proposed algorithm. Because of biological variations, a tissue

type may have inherent heterogeneity in it. A typical example is tumor, which are naturally heterogeneous. Activity concentration in a small tissue structure can be underestimated or overestimated due to partial volume effects, which cause the structure being imaged to mix with adjacent structures of possibly markedly different kinetics within the image volume, resulting in a mixed kinetics of the structures involved. As a finite number of clusters is assumed to be present in the raw PET data, the clustering algorithm will automatically look for the cluster centers that best represent the dataset without any *a priori* knowledge about the data and without violating the specified number of clusters. Therefore, certain regions which are indeterminate but their kinetics are similar, may be grouped together due to the constraint on the number of clusters, resulting in the formation of vague clusters. Further studies are required to investigate tissue heterogeneity in cluster analysis.

In earlier work, O'Sullivan [110] used cluster analysis as an intermediate step to extract "homogeneous" TACs from data containing a heterogeneous mix of kinetics resulting from spillover and partial volume effects for parametric mapping. This method is very similar to FADS but still there is a main difference between them: FADS extracts physiological factors (TACs) that could (theoretically) be found in the original data, whereas the set of "homogeneous" TACs does not necessarily correspond to the underlying physiology. In this current work, cluster analysis is used to extract kinetic data with different temporal characteristics as well as for parametric mapping. This is important for data analysis because data with different temporal behavior are better characterized by the extracted features seen in a spatial map. A spatial map is simpler to interpret when compared to the original multidimensional data. However, similar to O'Sullivan's approach [110], our method is data driven and is independent of the properties of tracer that may be required by other methods [111]. Thus, the clustering approach can be applied to a wide range of tracer studies.

It is anticipated that cluster analysis has a great deal of potential in PET data analysis for various neurodegenerative conditions (e.g. dementias) or diseases such as multiple system atrophy, Lewy Body disease, and Parkinsons disease where numerical values for glucose metabolism and the patterns of glucose hypometabolism may aid in the diagnosis and the assessment of disease progression. Localization of seizure foci in patients with refractory extratemporal epilepsy is also important as it is a difficult management problem for surgical epilepsy programs for this patient group. Functional segmentation may be

a useful tool in this regard. Investigation of the applicability of cluster analysis to whole-body PET for lesion localization and assessment of treatment response in a variety of oncological conditions will also be a fruitful research direction. Combining information provided by structural images for segmentation of functional image data will certainly become one of the key research areas in a new, hybrid PET/CT imaging technique, which will likely replace the existing PET alone facilities and will become a new standard of cancer imaging in the near future. With this hybrid imaging technique, precisely coregistered anatomical (CT) and functional (PET) images can be acquired in a single scanning session, and accurate localization of lesions could be achieved with the use of CT images as they provide very clear boundary delineation and anatomical information.

The above list of segmentation methods and applications are by no means complete. In fact, segmentation is one of the most difficult problems in medical image analysis but it is very useful in many medical imaging applications. Tremendous efforts have been devoted to cope with different segmentation problems. Continuing advances in exploitation and development of new conceptual approaches for effective extraction of all information and features contained in different types of multidimensional images are of increasingly importance in this regard. The following quote by the late philosopher, Sir Karl Popper, is worth noting when we think about new ideas and using analysis tools [133]:

> . . . at any moment we are prisoners caught in the framework of our theories; our expectations; our language. But . . . if we try, we can break out of our framework at any time. Admittedly, we shall find ourselves again in a framework, but it will be a better and roomier one; and we can at any moment break out of it again.

There is no magic method that suits all problems. One has to realize the strengths and limitations of the technique, and understand what kind of information the technique provides, and careful definition of the goals of segmentation is essential. It is also important to remember that new ideas and techniques may bring us something valuable that we are eager to see but something may be overlooked or missed out in the mean time, because we are bounded by the framework of the ideas or techniques, just like a prisoner, as Popper said. What we can only hope is that the new idea or the new technique, i.e. the prison, will be a better and roomier one where we can break out of it again at any time if there is a need!

Acknowledgment

This work was partially supported by the Hong Kong Polytechnic University under Grant G-YX13. Some of the results presented in this chapter were obtained in the period 1999 to 2002 during which the author was sustained financially by the National Health and Medical Research Council (NHMRC) of Australia.

Questions

1. *What is the major purpose of image segmentation? Why is it so significant in medical image analysis?*
2. *Identify the major classes of techniques for image segmentation.*
3. *List the advantages and disadvantage of using edge detection techniques for image segmentation.*
4. *What are the advantages and disadvantages of using manual region of interest (ROI) delineation with respect to using a template?*
5. *What are the disadvantages of using similarity mapping for segmentation?*
6. *What are the common and differences between principal component analysis (PCA) and factor analysis of dynamic structures (FADS)?*
7. *What are the major advantages of cluster analysis over other multivariate analysis approaches such as PCA and FA?*
8. *What are the major advantages of using clustering for characterizing tissue kinetics?*

Bibliography

[1] Rosenfeld, A. and Kak, A. C., Digital Image Processing, Academic Press, New York, 1982.

[2] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching, Comp. Vision Graph. Image Proc., Vol. 46, pp. 1–21, 1989.

[3] Lim, K. O. and Pfefferbaum, A., Segmentation of MR brain images into cerebrospinal fluid spaces, white, and gray matter, J. Comput. Assist. Tomogr., Vol. 13, pp. 588–593, 1989.

[4] Brzakovic, D., Luo, X. M., and Brzakovic, P., An approach to automated detection of tumors in mammograms, IEEE Trans. Med. Imaging, Vol. 9, pp. 233–241, 1990.

[5] Liang, Z., MacFall, J. R., and Harrington, D. P., Parameter estimation and tissue segmentation from multispectral MR images, IEEE Trans. Med. Imaging, Vol. 13, pp. 441–449, 1994.

[6] Ardekani, B. A., Braun, M., Hutton, B. F., Kanno, I., and Iida, H., A fully automatic multimodality image registration algorithm, J. Comput. Assist. Tomogr., Vol. 19, pp. 615–623, 1995.

[7] Bankman, I. N., Nizialek, T., Simon, I., Gatewood, O. B., Weinberg, I. N., and Brody, W. R., Segmentation algorithms for detecting microcalcifications in mammograms, IEEE Trans. Inform. Technol. Biomed., Vol. 1, pp. 141–149, 1997.

[8] Small, G. W., Stern, C. E., Mandelkern, M. A., Fairbanks, L. A., Min, C. A., and Guze, B. H., Reliability of drawing regions of interest for positron emission tomographic data, Psych. Res., Vol. 45, pp. 177–185, 1992.

[9] White, D. R., Houston, A. S., Sampson, W. F., and Wilkins, G. P., Intra- and interoperator variations in region-of-interest drawing and their effect on the measurement of glomerular filtration rates, Clin. Nucl. Med., Vol. 24, pp. 177–181, 1999.

[10] Hoffman, E. J., Huang, S. C., and Phelps, M. E., Quantitation in positron emission computed tomography, 1: Effect of object size, J. Comput. Assist. Tomogr., Vol. 3, pp. 299–308, 1979.

[11] Mazziotta, J. C., Phelps, M. E., Plummer, D., and Kuhl, D. E., Quantitation in positron emission compted tomography, 5: Physical-anatomical effects, J. Cereb. Blood Flow Metab., Vol. 5, pp. 734–743, 1981.

[12] Hutchins, G. D., Caraher, J. M., and Raylman, R. R., A region of interest strategy for minimizing resolution distortions in quantitative myocardial PET studies, J. Nucl. Med., Vol. 33, pp. 1243–1250, 1992.

[13] Welch, A., Smith, A. M., and Gullberg, G. T., An investigation of the effect of finite system resolution and photon noise on the bias and precision of dynamic cardiac SPECT parameters, Med. Phys., Vol. 22, pp. 1829–1836, 1995.

[14] Bezdek, J., Hall, L., and Clarke, L., Review of MR image segmentation techniques using pattern recognition, Med. Phys., Vol. 20, pp. 1033–1048, 1993.

[15] Mazziotta, J. C. and Koslow, S. H., Assessment of goals and obstacles in data acquisition and analysis from emission tomography: Report of a series of international workshops, J. Cereb. Blood Flow Metab., Vol. 7(Suppl. 1), pp. S1–S31, 1987.

[16] Mazziotta, J. C., Pelizzari, C. A., Chen, G. T., Bookstein, F. L., and Valentino, D., Region of interest issues: The relationship between structure and function in the brain, J. Cereb. Blood Flow Metab., Vol. 11, pp. A51–A56, 1991.

[17] Fu, K. S. and Mui, J. K., A survey on image segmentation, Pattern Recogn., Vol. 13, pp. 3–16, 1981.

[18] Haralick, R. M. and Shapiro, L. G., Survey: Image segmentation techniques, Comput. Vision Graphics Image Proc., Vol. 29, pp. 100–132, 1985.

[19] Pal, N. R. and Pal, S. K., A review on image segmentation techniques, Pattern Recogn., Vol. 26, pp. 1227–1249, 1993.

[20] Gonzalez, R. C. and Woods, R. E., Digital Image Processing, Addison-Wesley, Reading, MA, 1993.

[21] Castleman, K. R., Digital Image Processing, Prentice Hall, Upper Saddle River, NJ, 1996.

[22] Kittler, J., Illingworth, J., and Foglein, J., Threshold based on a simple image statistics, Comp. Vision Graph. Image Proc., Vol. 30, pp. 125–147, 1985.

[23] Chow, C. K. and Kaneko, T., Automatic boundary detection of the left ventricle from cineangiograms, Comput. Biomed. Res., Vol. 5, pp. 388–410, 1972.

[24] Marr, D. and Hildreth, E., Theory of edge detection, Proc. Roy. Soc. London, Vol. 27, pp. 187–217, 1980.

[25] Sun, Y., Lucariello, R. J., and Chiaramida, S. A., Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary arteriograms, IEEE Trans. Med. Imaging, Vol. 14, pp. 242–248, 1995.

[26] Faber, T. L., Akers, M. S., Peshock, R. M., and Corbett, J. R., Three-dimensional motion and perfusion quantification in gated single-photon emission computed tomograms, J. Nucl. Med., Vol. 32, pp. 2311–2317, 1991.

[27] Hough, P. V. C., A method and means for recognizing complex patterns, US Patent 3069654, 1962.

[28] Deans, S. R., The Radon Transform and Some of Its Applications, Wiley, New York, 1983.

[29] Radon, J., Über die bestimmung von funktionen durchihre integralwärte längs gewisser männigfaltigkeiten, Bertichte Säechsiche Akad. Wissenschaften (Leipzig), Math. Phys. Klass, Vol. 69, pp. 262–277, 1917.

[30] Kalviainen, H., Hirvonen, P., Xu, L., and Oja, E., Probabilistic and non-probabilistic Hough transforms: Overview and comparisons, Image Vision Comput., Vol. 13, pp. 239–252, 1995.

[31] Kassim, A., Tan, T., and Tan, K., A comparative study of efficient generalized Hough transforms techniques, Image Vision Comput., Vol. 17, pp. 737–748, 1999.

[32] Martelli, A., Edge detection using heuristic search methods, Comp. Graph. Image Proc., Vol. 1, pp. 169–182, 1972.

[33] Nilsson, N. J., Principles of Artificial Intelligence, Springer-Verlag, Berlin, 1982.

[34] Geiger, D., Gupta, A., Costa, A., and Vlontzos, J., Dynamic programming for detecting, tracking, and matching deformable contours, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 17, pp. 294–302, 1995.

[35] Barret, W. A. and Mortensen, E. N., Interactive live-wire boundary detection, Med. Image Analy., Vol. 1, pp. 331–341, 1996.

[36] Zucker, S., Region growing: Childhood and adolescence, Comp. Graph. Image Proc., Vol. 5, pp. 382–399, 1976.

[37] Hebert, T. J., Moore, W. H., Dhekne, R. D., and Ford, P. V., Design of an automated algorithm for labeling the cardiac blood pool in gated SPECT images of radiolabeled red blood cells, IEEE Trans. Nucl. Sci., Vol. 43, pp. 2299–2305, 1996.

[38] Kim, J., Feng, D. D., Cai, T. W., and Eberl, S., Automatic 3D temporal kinetics segmentation of dynamic emission tomography image using adaptive region growing cluster analysis, In: Proceedings of 2002 IEEE Medical Imaging Conference, Vol. 3, IEEE, Norfolk, VA, pp. 1580–1583, 2002.

[39] Hartigan, J. A., Clustering Algorithms, Wiley, New York, 1975.

[40] Cooper, L., M-dimensional location models: Application to cluster analysis, J. Reg. Sci., Vol. 13, pp. 41–54, 1973.

[41] Bezdek, J. C., Ehrlich, R., and Full, W., FCM: The fuzzy c-means clustering algorithm, Comp. Geosci., Vol. 10, pp. 191–203, 1984.

[42] Ball, G. H. and Hall, D. J., A clustering technique for summarizing multi-variate data, Behav. Sci., Vol. 12, pp. 153–155, 1967.

[43] Anderberg, M. R., Cluster Analysis for Applications, Academic Press, New York, 1973.

[44] McLachlan, G. J. and Krishnan, T., The EM Algorithm and Extensions, Wiley, New York, 1997.

[45] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput. Vis., Vol. 1, pp. 321–331, 1987.

[46] Terzopoulos, D. and Fleischer, K., Deformable models, Visual Comput., Vol. 4, pp. 306–331, 1988.

[47] Fischler, M. A. and Elschlager, R. A., The representation and matching of pictorial structures, IEEE Trans. Comput., Vol. 22, pp. 67–92, 1973.

[48] Widrow, B., The "rubber-mask" technique, Pattern Recogn., Vol. 5, pp. 175–211, 1973.

[49] McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A survey, Med. Image Analy., Vol. 1, pp. 91–108, 1996.

[50] Mykkänen, J. M., Tohka, J., and Ruotsalainen, U., Automated delineation of brain structures with snakes in PET, In: Physiological Imaging of the Brain with PET, Gjedde, A., Hansen, S. B., Knudsen, G., and Paulson, O. B., eds., Academic Press, San Diego, pp. 39–43, 2001.

[51] Chiao, P. C., Rogers, W. L., Fessler, J. A., Clinthorne, N. H., and Hero, A. O., Motion-based estimation with boundary side information or boundary regularization, IEEE Trans. Med. Imaging, Vol. 13, pp. 227–234, 1994.

[52] Chiao, P. C., Rogers, W. L., Clinthorne, N. H., Fessler, J. A., and Hero, A. O., Model-based estimation for dynamic cardiac studies using ECT, IEEE Trans. Med. Imaging, Vol. 13, pp. 217–226, 1994.

[53] Meltzer, C. C., Leal, J. P., Mayberg, H. S., Wagner, H. N., and Frost, J. J., Correction of PET data for partial volume effects in human cerebral cortex by MR imaging, J. Comput. Assist. Tomogr., Vol. 14, pp. 561–570, 1990.

[54] Müller-Gärtner, H. W., Links, J. M., Price, J. L., Bryan, R. N., McVeigh, E., Leal, J. P., Davatzikos, C., and Frost, J. J., Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects, J. Cereb. Blood Flow Metab., Vol. 12, pp. 571–583, 1992.

[55] Fox, P. T., Perlmutter, J. S., and Raichle, M. E., A stereotatic method of anatomical localization for positron emission tomography, J. Comput. Assist. Tomogr., Vol. 9, pp. 141–153, 1985.

[56] Talairach, J., Tournoux, P., and Rayport, M., Co-planar Stereotaxic Atlas of the Human Brain, Thieme, Inc., New York, 1988.

[57] Thompson, P. and Toga, A., A surface-based technique for warping three-dimensional images of the brain, IEEE Trans. Med. Imaging, Vol. 15, pp. 402–417, 1996.

[58] Bremner, J. D., Bronen, R. A., De Erasquin, G., Vermetten, E., Staib, L. H., Ng, C. K., Soufer, R., Charney, D. S., and Innis, R. B., Development and reliability of a method for using magnetic resonance imaging for the definition of regions of interest for positron emission tomography, Clin. Pos. Imag., Vol. 1, pp. 145–159, 1998.

[59] Maintz, J. B. A. and Viergever, M. A., A survey of medical image registration, Med. Imag. Analy., Vol. 2, pp. 1–37, 1998.

[60] Pelizzari, C. A., Chen, G. T. Y., Spelbring, D. R., Weichselbaum, R. R., and Chen, C. T., Accurate three-dimensional registration of CT, PET and/or MR images of the brain, J. Comput. Assist. Tomogr., Vol. 13, pp. 20–26, 1989.

[61] Woods, R. P., Mazziotta, J. C., and Cherry, S. R., MRI-PET registration with automated algorithm, J. Comput. Assisted Tomogr., Vol. 17, pp. 536–546, 1993.

[62] Rogowska, J., Similarity methods for dynamic image analysis, In: Proceedings of International AMSE Conference on Signals and Systems, Vol. 2, Warsaw, Poland, 15–17 July 1991, pp. 113–124.

[63] Barber, D. C., The use of principal components in the quantitative analysis of gamma camera dynamic studies, Phys. Med. Biol., Vol. 25, pp. 283–292, 1980.

[64] Rogowska, J. and Wolf, G. L., Temporal correlation images derived from sequential MR scans, J. Comput. Assist. Tomogr., Vol. 16, pp. 784–788, 1992.

[65] Bandettini, P. A., Jesmanowicz, A., Wong, E. C., and Hyed, J. S., Processing strategies for time-course datasets in functional MRI of the human brain, Magn. Res. Med., Vol. 30, pp. 161–173, 1993.

[66] Rogowska, J., Preston, K., Hunter, G. J., Hamberg, L. M., Kwong, K. K., Salonen, O., and Wolf, G. L., Applications of similarity mapping in dynamic MRI, IEEE Trans. Med. Imaging, Vol. 14, pp. 480–486, 1995.

[67] Jolliffe, I., Principal Component Analysis, Springer, New York, 1986.

[68] Pearson, K., On lines and planes of closest fit to systems of points in space, Phil. Mag., Vol. 6, pp. 559–572, 1901.

[69] Hotelling, H., Analysis of a complex of statistical variables into principal components, J. Edu. Psycho., Vol. 24, pp. 417–441, 1933.

[70] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., Numerical Recipes in C. The Art of Scientific Computing, Cambridge University Press, New York, 1992.

[71] Golub, G. H. and Van Loan, C. F., Matrix Computations, 3rd edn., John Hopkins University Press, Baltimore, 1996.

[72] Moeller, J. R. and Strother, S. C., A regional covariance approach to the analysis of functional patterns in positron emission tomographic data, J. Cereb. Blood Flow Metab., Vol. 11, pp. A121–A135, 1991.

[73] Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S., Functional connectivity: The principal component analysis of large (PET) data sets, J. Cereb. Blood Flow Metab., Vol. 13, pp. 5–14, 1993.

[74] Pedersen, F., Bergström, M., and Långström, B., Principal component analysis of dynamic positron emission tomography images, Eur. J. Nucl. Med., Vol. 21, pp. 1285–1292, 1994.

[75] Strother, S. C., Anderson, J. R., Schaper, K. A., Sidtis, J. S., and Rottenberg, D. A., Linear models of orthogonal subspaces and networks from functional activation PET studies of the human brain, In: Information Processing in Medical Imaging, Bizais, Y., Barillot, C., and Di Paola, R., eds., Kluwer, Dordrecht, The Netherlands, pp. 299–310, 1995.

[76] Ardekani, B. A., Strother, S. C., Anderson, J. R., Law, I., Paulson, O. B., Kanno, I., and Rottenberg, D. A., On the detection of activation patterns using principal components analysis, In: Quantitative Functional Brain Imaging with Positron Emission Tomography, Carson, R. E., Daube-Witherspoon, M. E., and Herscovitch, P., eds., Academic Press, San Diego, pp. 253–257, 1998.

[77] Anzai, Y., Minoshima, S., Wolf, G. T., and Wahl, R. L., Head and neck cancer: Detection of recurrence with three-dimensional principal components analysis at dynamic FDG PET, Radiology, Vol. 212, pp. 285–290, 1999.

[78] Andersen, A. H., Gash, D. M., and Avison, M. J., Principal component analysis of the dynamic response measured by fMRI: A generalized linear systems framework, Mag. Res. Imag., Vol. 17, pp. 795–815, 1999.

[79] Baumgartner, R., Ryner, L., Richter, W., Summers, R., Jarmasz, M., and Somorjai, R., Comparison of two exploratory data analysis methods for fMRI: Fuzzy clustering vs. principal component analysis, Mag. Res. Imag., Vol. 18, pp. 89–94, 2000.

[80] Correia, J., A bloody future for clinical PET? [editorial], J. Nucl. Med., Vol. 33, pp. 620–622, 1992.

[81] Iida, H., Rhodes, C. G., De Silva, R., Araujo, L. I., Bloomfield, P. M., Lammertsma, A. A., and Jones, T., Use of the left ventricular time-activity curve as a non-invasive input function in dynamic Oxygen-15-Water positron emission tomography, J. Nucl. Med., Vol. 33, pp. 1669–1677, 1992.

[82] Chen, K., Bandy, D., Reiman, E., Huang, S. C., Lawson, M., Feng, D., Yun, L. S., and Palant, A., Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, ^{18}F-fluoro-2-deoxyglucose, the Patlak method, and an image-derived input function, J. Cereb. Blood Flow Metab., Vol. 18, pp. 716–723, 1998.

[83] Houston, A. S., The effect of apex-finding errors on factor images obtained from factor analysis and oblique transformation, Phys. Med. Biol., Vol. 29, pp. 1109–116, 1984.

[84] Nirjan, K. S. and Barber, D. C., Factor analysis of dynamic function studies using *a priori* physiological information, Phys. Med. Biol., Vol. 31, pp. 1107–1117, 1986.

[85] Šámal, M., Kárný, M., Sùrová, H., and Dienstbier, Z., Rotation to simple structure in factor analysis of dynamic radionuclide studies, Phys. Med. Biol., Vol. 32, pp. 371–382, 1987.

[86] Buvat, I., Benali, H., Frouin, F., Bazin, J. P., and Di Paola, R., Target apex-seeking in factor analysis on medical sequences, Phys. Med. Biol., Vol. 38, pp. 123–128, 1993.

[87] Sitek, A., Di Bella, E. V. R., and Gullberg, G. T., Factor analysis with *a priori* knowledge—Application in dynamic cardiac SPECT, Phys. Med. Biol., Vol. 45, pp. 2619–2638, 2000.

[88] Wu, H. M., Hoh, C. K., Buxton, D. B., Schelbert, H. R., Choi, Y., Hawkins, R. A., Phelps, M. E., and Huang, S. C., Factor analysis for extraction of blood time-activity curves in dynamic FDG-PET studies, J. Nucl. Med., Vol. 36, pp. 1714–1722, 1995.

[89] Wu, H. M., Huang, S. C., Allada, V., Wolfenden, P. J., Schelbert, H. R., Phelps, M. E., and Hoh, C. K., Derivation of input function from FDG-PET studies in small hearts, J. Nucl. Med., Vol. 37, pp. 1717–1722, 1996.

[90] Sitek, A., Di Bella, E. V. R., and Gullberg, G. T., Factor analysis of dynamic structures in dynamic SPECT imaging using maximum entropy, IEEE Trans. Nucl. Sci., Vol. 46, pp. 2227–2232, 1999.

[91] Sitek, A., Gullberg, G. T., and Huesman, R. H., Correction for ambiguous solutions in factor analysis using a penalized least squares objective, IEEE Trans. Med. Imaging, Vol. 21, pp. 2166–225, 2002.

[92] Ashburner, J., Haslam, J., Taylor, C., Cunningham, V. J., and Jones, T., A cluster analysis approach for the characterization of dynamic PET data, In: Quantification of Brain Function using PET, Myers, R., Cunningham, V., Bailey, D., and Jones, T., eds., Academic Press, San Diego, pp. 301–306, 1996.

[93] Acton, P. D., Pilowsky, L. S., Costa, D. C., and Ell, P. J., Multivariate cluster analysis of dynamic iodine-123 iodobenzamide SPET dopamine D_2 receptor images in schizophrenia, Eur. J. Nucl. Med., Vol. 24, pp. 111–118, 1997.

[94] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Segmentation of dynamic PET images using cluster analysis, IEEE Trans. Nucl. Sci., Vol. 49, pp. 200–207, 2002.

[95] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Simultaneous estimation of physiological parameters and the input function—*In vivo* PET data, IEEE Trans. Inform. Technol. Biomed., Vol. 5, pp. 67–76, 2001.

[96] Wong, K. P., Meikle, S. R., Feng, D., and Fulham, M. J., Estimation of input function and kinetic parameters using simulated annealing: Application in a flow model, IEEE Trans. Nucl. Sci., Vol. 49, pp. 707–713, 2002.

[97] Cunningham, V. J. and Jones, T., Spectral analysis of dynamic PET studies, J. Cereb. Blood Flow Metab., Vol. 13, pp. 15–23, 1993.

[98] Zubal, I. G., Harrell, C. R., Smith, E. O., Rattner, Z., Gindi, G., and Hoffer, P. B., Computerized three-dimensional segmented human anatomy, Med. Phys., Vol. 21, pp. 299–302, 1994.

[99] Hoffman, E. J., Cutler, P. D., Digby, W. M., and Mazziotta, J. C., 3-D phantom to simulate cerebral blood flow and metabolic images for PET, IEEE Trans. Nucl. Sci., Vol. 37, pp. 616–620, 1990.

[100] Hawkins, R. A., Phelps, M. E., and Huang, S. C., Effects of temporal sampling, glucose metabolic rates, and disruptions of the blood-brain barrier on the FDG model with and without a vascular compartment: Studies in human brain tumors with PET, J. Cereb. Blood Flow Metab., Vol. 6, pp. 170–183, 1986.

[101] Akaike, H., A new look at the statistical model identification, IEEE Trans. Automatic Control, Vol. AC-19, pp. 716–723, 1974.

[102] Schwarz, G., Estimating the dimension of a model, Ann. Stat., Vol. 6, pp. 461–464, 1978.

[103] Hooper, P. K., Meikle, S. R., Eberl, S., and Fulham, M. J., Validation of post injection transmission measurements for attenuation correction in neurologic FDG PET studies, J. Nucl. Med., Vol. 37, pp. 128–136, 1996.

[104] Huang, S. C., Phelps, M. E., Hoffman, E. J., Sideris, K., Selin, C., and Kuhl, D. E., Noninvasive determination of local cerebral metabolic rate of glucose in man, Am. J. Physiol., Vol. 238, pp. E69–E82, 1980.

[105] Patlak, C. S., Blasberg, R. G., and Fenstermacher, J., Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data, J. Cereb. Blood Flow Metab., Vol. 3, pp. 1–7, 1983.

[106] Gunn, R. N., Lammertsma, A. A., and Cunningham, V. J., Parametric imaging of ligand-receptor interactions using a reference tissue model and cluster analysis, In: Quantitative Functional Brain Imaging with Positron Emission Tomography, Carson, R. E., Daube-Witherspoon, M. E., and Herscovitch, P., eds., Academic Press, San Diego, pp. 401–406, 1998.

[107] Lammertsma, A. A. and Hume, S. P., Simplified reference tissue model for PET receptor studies, NeuroImage, Vol. 4, pp. 153–158, 1996.

[108] Gunn, R. N., Lammertsma, A. A., Hume, S. P., and Cunningham, V. J., Parametric imaging of ligand-receptor binding in PET using a simplified reference region model, NeuroImage, Vol. 6, pp. 279–287, 1997.

[109] Wong, K. P., Feng, D., Meikle, S. R., and Fulham, M. J., Non-invasive determination of the input function in PET by a Monte Carlo approach and cluster analysis, J. Nucl. Med., Vol. 42, No. 5(Suppl.), p. 183P, 2001.

[110] O'Sullivan, F., Imaging radiotracer model parameters in PET: A mixture analysis approach, IEEE Trans. Med. Imaging, Vol. 12, pp. 399–412, 1993.

[111] Kimura, Y., Hsu, H., Toyama, H., Senda, M., and Alpert, N. M., Improved signal-to-noise ratio in parametric images by cluster analysis, NeuroImage, Vol. 9, pp. 554–561, 1999.

[112] Bentourkia, M., A flexible image segmentation prior to parametric estimation, Comput. Med. Imaging Graphics, Vol. 25, pp. 501–506, 2001.

[113] Kimura, Y., Senda, M., and Alpert, N. M., Fast formation of statistically reliable FDG parametric images based on clustering and principal components, Phys. Med. Biol., Vol. 47, pp. 455–468, 2002.

[114] Zhou, Y., Huang, S. C., Bergsneider, M., and Wong, D. F., Improved parametric image generation using spatial-temporal analysis of dynamic PET studies, NeuroImage, Vol. 15, pp. 697–707, 2002.

[115] Bal, H., DiBella, E. V. R., and Gullberg, G. T., Parametric image formation using clustering for dynamic cardiac SPECT, IEEE Trans. Nucl. Sci., Vol. 50, pp. 1584–1589, 2003.

[116] Toyama, H., Takazawa, K., Nariai, T., Uemura, K., and Senda, M., Visualization of correlated hemodynamic and metabolic functions in cerebrovascular disease by a cluster analysis with PET study, In: Physiological Imaging of the Brain with PET, Gjedde, A., Hansen, S. B., Knudsen, G. M., and Paulson, O. B., eds., Academic Press, San Diego, pp. 301–304, 2001.

[117] Koh, W. J., Rasey, J. S., Evans, M. L., Grierson, J. R., Lewellen, T. K., Graham, M. M., Krohn, K. A., and Griffin, T. W., Imaging of hypoxia in human tumors with [F-18]fluoromisonidazole, Int. J. Radiat. Oncol. Biol. Phys., Vol. 22, pp. 199–212, 1992.

[118] Marsden, P. K., Personal communication, 2003.

[119] Huang, S. C., Hoffman, E. J., Phelps, M. E., and Kuhl, D. E., Quantitation in positron emission computed tomography, 2: Effects of inaccurate attenuation correction, J. Comput. Assist. Tomogr., Vol. 3, pp. 804–814, 1979.

[120] Dahlbom, M. and Hoffman, E. J., Problems in signal-to-noise ratio for attenuation correction in high-resolution PET, IEEE Trans. Nucl. Sci., Vol. 34, pp. 288–293, 1987.

[121] Huang, S. C., Carson, R. E., Phelps, M. E., Hoffman, E. J., Schelbert, H. R., and Kuhl, D. E., A boundary method for attenuation correction in positron computed tomography, J. Nucl. Med., Vol. 22, pp. 627–637, 1981.

[122] Xu, M., Luk, W. K., Cutler, P. D., and Digby, W. M., Local threshold for segmented attenuation correction of PET imaging of the thorax, IEEE Trans. Nucl. Sci., Vol. 41, pp. 1532–1537, 1994.

[123] Meikle, S. R., Dahlbom, M., and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography, J. Nucl. Med., Vol. 34, pp. 143–144, 1993.

[124] Papenfuss, A. T., O'Keefe, G. J., and Scott, A. M., Segmented attenuation correction in whole body PET using neighbourhood EM clustering, In: 2000 IEEE Medical Imaging conference, IEEE Publication, Lyon, France, 2000.

[125] Bettinardi, V., Pagani, E., Gilardi, M. C., Landoni, C., Riddell, C., Rizzo, G., Castiglioni, I., Belluzzo, D., Lucignani, G., Schubert, S., and Fiazio, F., An automatic classification technique for attenuation correction in positron emission tomography, Eur. J. Nucl. Med., Vol. 26, pp. 447–458, 1999.

[126] Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W., Brain magnetic resonance imaging with contrast dependent on blood oxygenation, Proc. Natl. Acad. Sci. USA, Vol. 87, pp. 9868–9872, 1990.

[127] Bullmore, E. and Brammer, B., Statistical methods of estimation and inference for functional MR image analysis, Magn. Reson. Med., Vol. 35, pp. 261–277, 1996.

[128] Lange, N., Statistical approaches to human brain mapping by functional magnetic resonance imaging, Stat. Med., Vol. 15, pp. 389–428, 1996.

[129] Moser, E., Diemling, M., and Baumgartner, R., Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part II: Quantification, J. Magn. Reson. Imaging, Vol. 7, pp. 1102–1108, 1997.

[130] Goutte, C., Toft, P., Rostrup, E., Nielsen, F. Å., and Hansen, L. K., On clustering fMRI time series, NeuroImage, Vol. 9, pp. 298–310, 1999.

[131] Fadili, M. J., Ruan, S., Bloyet, D., and Mazoyer, B., A multistep unsupervised fuzzy clustering analysis of fMRI time series, Hum. Brain Mapping, Vol. 10, pp. 160–178, 2000.

[132] Schmidt, K., Lucignani, G., Moresco, R. M., Rizzo, G., Gilardi, M. C., Messa, C., Colombo, F., Fazio, F., and Sokoloff, L., Errors introduced by tissue heterogeneity in estimation of local cerebral glucose utilization with current kinetic models of the [^{18}F]fluorodeoxyglucose method, J. Cereb. Blood Flow Metab., Vol. 12, pp. 823–834, 1992.

[133] Popper, K. R., Normal science and its dangers, In: Criticism and the Growth of Knowledge, Lakatos, I. and Musgrave, A., eds., Cambridge University Press, Cambridge, pp. 51–58, 1970.

Chapter 4

Automatic Segmentation of Pancreatic Tumors in Computed Tomography

Maria Kallergi,[1] *Marla R. Hersh*,[1] *and Anand Manohar*[1]

4.1 Introduction

Pancreatic cancer is the fourth leading cause of cancer deaths in the United States but only the tenth site for new cancer cases (estimated at 30,300 in 2002) [1, 2]. The reason for this major difference is that there are no clear early symptoms of pancreatic cancer and no screening procedures or screening policy for this disease. It is usually diagnosed at a late stage and has a poor prognosis with a 1-year survival rate of 20% and a 5-year survival rate of less than 5% [3]. Complete surgical resection is the only way to significantly improve prognosis and possibly lead to a cure. Unfortunately, only 15–20% of the patients can undergo resection with median survival rates from 12 to 19 months and a 5-year survival rate of 15–20%. A large majority of patients with pancreatic cancer receives palliative care or may follow a therapeutic approach the impact of which is currently quite limited and difficult to assess quantitatively [3].

Pancreatic cancer is a disease that is not extensively studied and is poorly understood. A significant risk factor associated with pancreatic cancers is age (frequency increases linearly after 50 years of age) with median age of diagnosis of 71 years. In addition to aging, other probable risk factors include family history, cigarette smoking, long-standing diabetes, and chronic and hereditary pancreatitis [2]. Studies have also implicated, without any consistency,

[1]Department of Radiology, H. Lee Moffitt Cancer Center & Research Institute, University of South Florida, Tampa, FL 33612

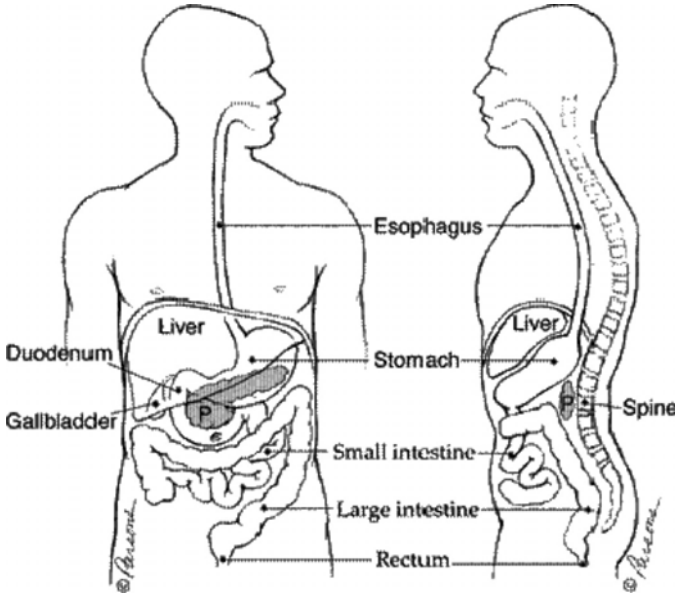


Figure 4.1: Human anatomy indicating the location of the pancreas and major neighboring organs (reprinted from http://pathology2.jhu.edu/pancreas/pancreas1.cfm).

a number of other factors including alcohol consumption, diet and nutrition, occupational exposures, genetic predisposition, multiple endocrine neoplasia, hereditary nonpolyposis colorectal cancer, familial adenomatous polyposis and Gardner syndrome, familial atypical multiple mole melanoma syndrome, von Hippel-Lindau syndrome, and germline mutations in the *BRCA2* gene [4–6]. Tumor markers have been used for pancreatic cancer including carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA 19-9) but no conclusive results exist to date [2].

Figure 4.1 shows a drawing of the human anatomy, the location of the pancreas, and major surrounding organs and structures. The pancreas lies in the upper abdomen, transversely behind the stomach, with its tip very close to the hilum of the spleen. It is an elongated organ with three major parts: head, body, and tail as shown in Fig. 4.2. The pancreas is a gland that produces regulating hormones and enzymes for protein breakdown. The enzymes are secreted into the duodenum through a set of tubes called pancreatic ducts.

Pancreatic cancer mostly originates in the ductal system of the pancreas (95% of the cases) and is termed ductal adenocarcinoma [7]; pancreatic cancer that affects the endocrine cells is called islet cell cancer [8]. About 60% of all

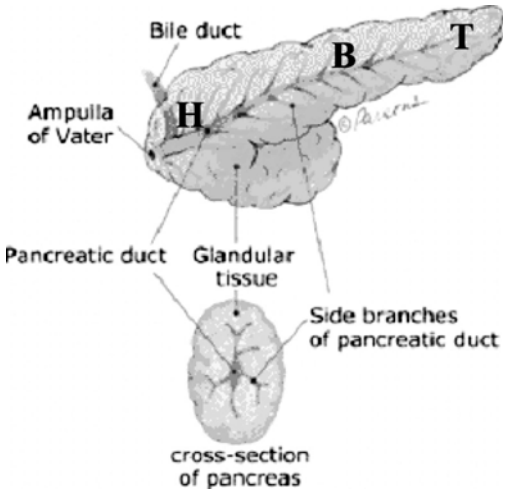


Figure 4.2: Detailed structure and main parts of the pancreas where **H** = head, **B** = body, and **T** = tail (reprinted from http://pathology2.jhu.edu/pancreas/pancreas1.cfm).

pancreatic cancers occur in the head or neck of the pancreas, about 15% occur in the body of the pancreas, about 5% in the tail, and about 20% are diffused through the gland [7]. Pancreatic cancer metastasizes rapidly even when the primary tumors are less than 2 cm. Metastasis most commonly occurs to regional lymph nodes, then to liver, and less commonly to the lungs. The tumors display a high degree of resistance to conventional chemotherapy and radiation therapy.

The National Cancer Institute (NCI) created a Review Group in 2001 to define an agenda for action for pancreatic cancer and the research priorities that could reduce morbidity and mortality from this difficult disease [2]. The research priorities set by NCI spanned a wide range of areas including

- tumor biology
- risk, prevention, detection, and diagnosis
- therapy
- health sciences research

In each of the areas above, research priorities were defined that could lead to the advancement of our knowledge of the disease and improved health care for the patients. Our interest in this chapter is on the role of imaging and computer

techniques on the detection, diagnosis, and surveillance of pancreatic cancer [9]. Hence, this chapter reviews the imaging priorities set by the Review Group [2], the currently used imaging techniques and state-of-the-art computer methodologies, the requirements for development of computer algorithms for the automatic diagnosis, management, and treatment of the disease, and the challenges posed to the development of computer methodologies. A novel algorithm is also described for the segmentation and classification of pancreatic tumor. Preliminary results are reported from its application on computed tomography (CT) images of the pancreas.

4.2 Imaging of the Pancreas

4.2.1 Imaging Modalities

Imaging of the pancreas is done for diagnosis, staging, and surveillance of benign or malignant conditions of the gland; surveillance includes monitoring of the disease, the effects of treatments, and the biological processes [10]. Imaging is faced with formidable obstacles in the case of pancreatic cancer because of the complexity of the disease and its high incidence of metastasis. Early neoplastic changes are difficult to detect and diagnose. Most of the available imaging techniques fail to detect early signs of pancreatic cancer. It is almost always detected after it has been spread beyond the gland. Imaging of the pancreas may be done in a variety of ways:

1. Computed tomography (CT); standard [11, 12], helical [7, 13, 14], and multidetector [8, 15] with the latter becoming the dominant modality of choice.
2. Magnetic resonance imaging (MRI) [9, 16, 17].
3. Magnetic resonance cholangiopancreatography (MRCP) [18, 19].
4. Endoscopic retrograde cholangiopancreatography (ERCP) [20]
5. Endoscopic ultrasonography (EUS) [21, 22]
6. Positron emission tomography (PET) [23, 24]

The usage of one technique over another depends on availability, purpose of imaging, and expertise. Additional considerations in the selection of an imaging

modality include the desired accuracy of the procedure for providing staging information, or its ability to perform simultaneous biopsy of the tumor, or its capacity to facilitate therapeutic procedures. Detection usually starts with transabdominal sonography to identify causes of pain. After sonography, CT is used as the primary modality for diagnosis and staging. MRI is also used for staging. MRCP and ERCP imaging provide additional information on the level of obstruction of the biliary or pancreatic ductal systems. Fine-needle aspiration of suspected pancreatic lesions can be done with EUS for increased biopsy specificity. Specificity is a problem with all imaging modalities as they do not make it possible to distinguish between pancreatic cancer and other pancreatic pathology, e.g., chronic pancreatitis, mucinous cystadenoma, and intraductal papillary mucinous neoplasms [10].

Today the most common modality for pancreatic imaging is helical CT, which has significantly improved outcomes relative to the standard CT or the other imaging modalities. Standard abdominal CT scans can help detect 70–80% of pancreatic carcinomas [3, 13]. But 40–50% of tumors smaller than 3 cm are missed, and these are the tumors most likely to be resectable. Helical CT improved significantly the resolution of conventional CT for pancreatic tumor imaging [25]. Helical CT has also impacted staging and treatment monitoring procedures and is now probably the most useful imaging technique for such investigations. Helical CT is the technique that will be focused on in this chapter.

Priorities set by the NCI Review Group for pancreatic cancer imaging include [2] (a) increase specificity of current imaging modalities, (b) increase sensitivity of current imaging modalities for small invasive and preinvasive lesions in both normal and abnormal pancreas, (c) develop and test molecular imaging techniques, (e) develop and test screening and surveillance imaging protocols for high-risk patients, and (f) develop and test noninvasive techniques to accurately define the effect of treatment. Computer aided diagnosis (CAD) schemes are computer algorithms that could assist the physician (radiologist or oncologist) in the interpretation of the radiographic images and the evaluation of the disease. CAD could play several roles in the above imaging priorities and contribute in several recommendations and research directions. CAD using CT scans seems to be the logical first step in the development of computer tools for pancreatic cancer because of the major role of CT in this area, the large amount of information available in CT scans, and the considerable potential for improvements that could have significant clinical impact independent of magnitude.

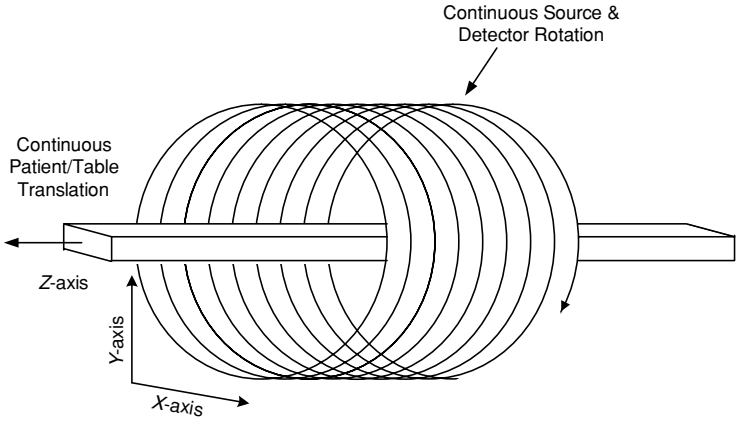


Figure 4.3: Schematic diagram of the helical CT set-up and operation principle including cross-section (slice) plane (x,y), and z axis orientation.

4.2.2 Helical CT Imaging Characteristics

CT was developed in the early 1960s with the first clinical system installed in 1971 [26]. Today, there are five generations of CT scanners characterized by different scanning conditions and properties. Helical or spiral CT is the latest generation of scanners that combines a continuous rotation of the X-ray source and ring of detectors with a continuous movement of the examination table. Hence, data are acquired continuously while the patient is moved through the gantry [27, 28]. Figure 4.3 shows a drawing of the helical CT scanning principle and associated orientations [7]. Figure 4.4 shows a typical two-dimensional (2-D) CT slice through the abdomen. The major organs and structures are labeled including a tumor on the head of the pancreas (Arrow B).

Basic CT principles and operation description can be found in a variety of journal articles and dedicated books [27–29]; Van Hoe and Baert [7] provide a comprehensive summary of the advantages of helical CT relative to standard CT scans for pancreatic tumors, including a description of the various imaging parameters. Kopecky and colleagues from Indiana University Medical Center provide an excellent illustrative presentation of the physical characteristics of helical CT at http://www.indyrad.iupui.edu/public/lectures/multislice/sld001.htm. Herein we focus on the helical CT imaging parameters that may be of interest to the medical imaging scientist and engineer and pertinent to computer applications and CAD development. These parameters include the following:

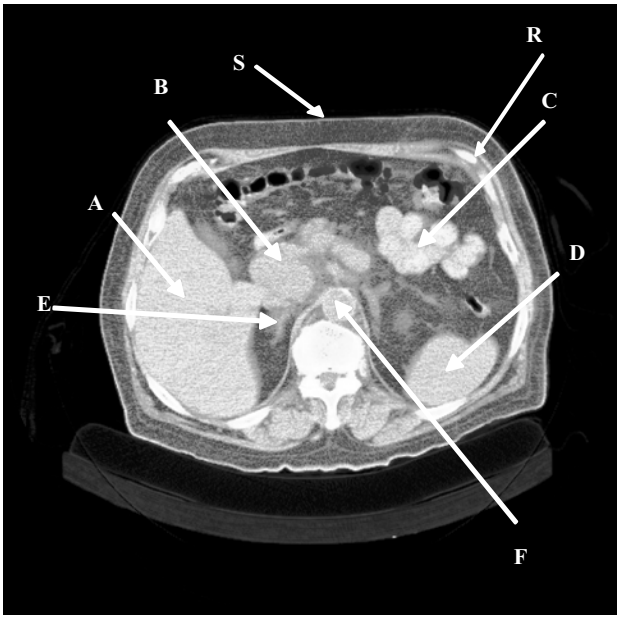


Figure 4.4: Contrast enhanced, helical CT scan through the abdomen and the head of the pancreas obtained with a reconstruction width of 8 mm (equal to slice thickness). A = liver; B = head of pancreas with tumor; C = bowel; D = spleen; E = right adrenal; F = aorta.

(a) *Slice thickness (mm)*: This is equal to the collimation of the X-ray beam. Maximum slice thickness depends on the size of the detectors and is typically about 8–10 mm. Thick slices are used for general abdominal scans and they usually have better contrast resolution than do thin slices, which in turn have better spatial resolution and require higher radiation dose. Thin slices are used for small organs or to evaluate and review a region of interest in more detail. Pancreatic scans may be performed in either thick (8 mm) scans or thin (less than 4 mm) scans or combinations depending on the case and/or the protocol requirements. Different slice thicknesses result in different image properties and noise characteristics so they have to be carefully considered in computer applications including registration, reconstruction, and segmentation.

(b) *Incrementation*: This is related to the longitudinal or z-axis resolution and is also referred to as step or index. It defines the distance between consecutive slices. It is usually equal to the slice thickness leaving no gaps

between consecutive slices. In special cases, e.g., small organs or three-dimensional (3-D) reconstructions, it may be smaller than slice thickness yielding an overlap between consecutive slices. Incrementation may be changed retrospectively in helical CT.

(c) *Resolution*: The field of view is related to the spatial resolution of the CT slices or the in-plane resolution. Based on the field of view and the size of the 2-D matrix (X and Y dimensions in pixels (see Fig. 4.3)) the spatial resolution or pixel size can be determined. The dynamic resolution or pixel depth is determined by the characteristics of the detector.

(d) *Exposure*: kVp and mAs are the two parameters that influence exposure with kVp defining the beam quality or the average intensity and mAs defining the quantity of the beam. For larger patients, the mAs may be increased.

(e) *Pitch*: This is the ratio of the distance traveled by the table during one full rotation of the detector (gantry) to the beam collimation. For example, a pitch of 1.0 corresponds to a scan with a beam collimation of 10 mm, 1-sec duration of a 360 degree rotation, and a rate of table movement of 10 mm/sec. By doubling the rate of table translation, the pitch is increased proportionally. In pancreatic screening scans the table speed is on the order of 15 mm/rotation and the pitch is 6. In diagnostic high-resolution scans the table speed is on the order of 6 with a pitch of 6. The higher the pitch the shorter the scan time. For multislice helical CT, *pitch* is also defined as the ratio of the table travel per gantry rotation to the nominal slice thickness. This is a more ambiguous definition, not applicable to single-slice helical CT, but often used by the manufacturers of multislice CT scanners [30].

(f) *Contrast type and amount*: For pancreatic scans, a contrast material (e.g., Omnipaque 320 or 350) is administered intravenously prior to imaging at a volume of 100–120 cc. Water, gastograffin, or barium is usually given as oral contrast. There is 50–60 sec scan delay to allow for optimum imaging of the pancreas after the administration of the contrast material.

(g) *Data reconstruction interval*: This is the thickness of the reconstructed slices and is usually equal to the scanned slice thickness.

Three-dimensional reconstructed CT volumes may be used to generate 2-D views of the organ in the coronal (*XZ* plane) and sagittal (*YZ* plane) modes in additional to the traditional transaxial view (*XY* plane) (Fig. 4.3).

A standard CT scan of the abdomen and pelvis consists of about 40 slices. An abdominal high-resolution scan of the pancreas may consist of 5–40 slices depending on the selected slice thickness and the orientation of the gland [31]. Slice thickness, slice interval, and slice starting point may be selected retrospectively in helical CT scans because of the continuous nature of the process. This is an advantage of the helical over the conventional CT imaging where these parameters need to be determined at the beginning of the scan at the risk of improper imaging of a lesion [32].

Despite its many advantages, e.g., high-spatial resolution and ability to identify vascular involvement, helical CT fails to detect small (<2cm) tumors, hepatic metastasis, and peritoneal implants, and also has low specificity to other pancreatic pathology [33]. Most patients with pancreatic cancer are evaluated by combinations of imaging tests rather than a single test. In addition to the limitations of the imaging technique, pancreatic tumor measurements and evaluation during treatment monitoring are done visually using subjective criteria by a single expert or, at best, a panel of experts. Hence, there is significant interobserver, and possibly intraobserver, variability in the evaluation of response to treatment and management of the patients with pancreatic cancer.

Important to the development of image processing techniques for pancreatic cancer applications is the knowledge of the clinical imaging characteristics of the normal and abnormal pancreas. The normal pancreas is relatively easy to delineate on CT slices. Understanding how the image of the normal pancreas may be distorted by disease and particularly pancreatic masses (benign or malignant) is the basis for selecting robust features for the development of automated segmentation, classification, registration, and reconstruction methodologies. The most important features used by the radiologists and oncologists in the evaluation of pancreatic adenocarcinoma on radiologic images are summarized in Table 4.1. These features are merely general observation that may not always hold.

There is variability in the imaging characteristics of pancreatic tumors on CT images that increases detection and diagnostic difficulty relative to other organ abnormalities. Table 4.1 summarizes the characteristics of adenocarcinoma

Table 4.1: Helical CT imaging characteristics of the normal pancreas and changes induced by pancreatic adenocarcinoma [2, 31, 34–36]

Normal pancreas	Abnormal pancreas (Adenocarcinoma)
• Uniform density (image intensity) throughout, slightly lower than that of the liver, no calcifications. Contrast material increases pancreatic density uniformly. Approximately equal density to the spleen, kidneys, and skeletal muscle.	• Variable imaging characteristics; tumors generally appear isodense to normal pancreatic tissue in enhanced studies. Some adenocarcinomas may show central necrosis or appear as hyperdense areas relative to the rest of the pancreas.
• The contour is smooth with a faint lobulation is some cases.	• Abrupt transition to the smooth contour may occur due to the presence of a mass.
• A fat plane usually surrounds the normal pancreas with the exception of very thin patients. The fat plane appears as an area of lower intensity than the gland area on the CT scan.	• The fat plane is usually disrupted or disappears due the presence of a mass or other disease.
• The anterior–posterior diameter of the normal pancreas averages 3 cm in the head, 2.5 cm in the body, and 2 cm in the tail.	• Changes in the size of the pancreas may occur due to the presence of large masses.
• Organ tends to taper uniformly from head to the tail.	• Duct dilation is one of the most significant consequence of pancreatic adenocarcinomas.
• Reports suggest that the ratio between the transverse diameter of the accompanying vertebral body and the pancreas can be used as a guide for normalcy.	• Alterations occur in organs and structures adjacent to the pancreas due to the presence of masses.
• There is usually a fatty appearance due to the gland's nature.	• Gland areas are enhanced with contrast material that could allow separation from normal tissues. Usually appear as hypodense areas due to poor arterial blood supply.

because it is the most common type of pancreatic cancer and, hence, the one better understood. Similar imaging and evaluation procedures are initially followed for all pancreatic tumor types.

All pancreatic tumors are better visualized when intravenous contrast material is used. Only necrotic tumors and very large tumors can be identified without contrast enhancement. Endocrine tumors often have associated calcifications and are less likely to have central necrosis than do adenocarcinomas.

They also enhance more than normal tissue during the initial phases of contrast administration [36]. Cystic neoplasms have a variety of appearances. They can appear solid secondary to the multiple tiny nonvisible cysts or they can appear as multiple small cysts or as "multilocular-appearing mass" with thin septations [37]. Alterations in the bowel, blood vessels, or ducts within or adjacent to the pancreas may be caused by all types of pancreatic tumors and are important features in the identification of pancreatic abnormalities [38].

Once diagnosed, pancreatic tumors are surgically removed or treated. The resection of pancreatic tumors is based on the identified tumor size and the presence or absence of additional abnormal signs on the abdominal CT scans. Resection is determined by three imaging criteria:

- Tumor size (less than 4 cm usually); tumors greater than 5 cm are resectable in less than 10% of the cases.
- Vascular invasion, in particular invasion of superior mesenteric artery/vein or portal vein.
- Presence of malignant ascites, nodal disease outside of the area of resection, liver metastases, or peritoneal carcinomatosis.

Presence of metastatic disease, involvement of the mesenteric, and invasion of the portal or superior mesenteric vein are all indicators of nonresectable disease [3].

In addition to the imaging characteristics of pancreatic cancer, clinical findings contribute to the diagnosis and management of the disease. Clinical and demographic characteristics that may be useful in CAD development include [3]

- age; one of the most significant risk factors for pancreatic cancer.
- presence of jaundice that is usually associated with adenocarcinoma of the pancreatic head; resectability rate of pancreatic tumors is noted to be higher in these patients than in patients not presenting with jaundice.
- abdominal pain that may be used as a survival predictor; shorter survival intervals are associated with greater pain reported prior to surgery.
- weight loss and anorexia symptoms.
- diabetes onset.

Clinical and demographic characteristics play a role in feature selection for clustering and classification. In the past, few CAD application incorporated image and nonimage characteristics in algorithm design. New directions in medical image analysis and processing clearly demonstrate the need to consider the patient as a whole and integrate information from a variety of sources to achieve high performances.

4.3 Computer Applications in Pancreatic Cancer Imaging

There is limited development of automatic approaches for the detection and/or diagnosis of pancreatic cancer either from CT or other imaging modalities. This is certainly an area worthy of further investigation and an area identified as in great need of technological advances by the NCI Review Group [2]. Imaging priorities set by the Group have been summarized earlier in this chapter. One of the most interesting recommendation was for a collaborative research and training approach that will link molecular biology, pathology, and imaging as well as for a well documented source of images to support computer applications and image processing [2].

A few common stages may be identified in all algorithms designed for medical imaging applications, including those designed for assisting the interpretation of CT scans. Figure 4.5 presents the basic modules of an algorithm that aims at assisting the physicians in the interpretation of CT images for the detection, diagnosis, and surveillance of disease. Registration and 3-D reconstruction may precede or follow the last stage of "Processing" (shown in Fig. 4.5) depending on

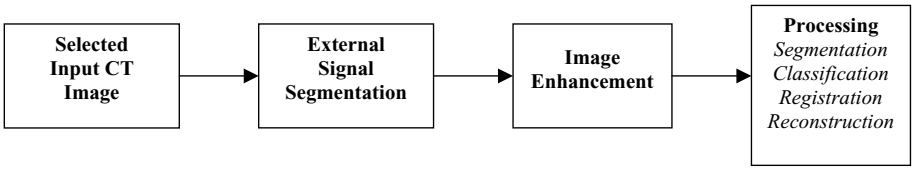


Figure 4.5: General algorithm design for CT image processing. Processing may include a segmentation, a classification, a registration, a reconstruction step, or any combinations of these.

the goals of the development. Herein we focus on issues related to 2-D CT processing and, hence, registration and reconstruction will not be discussed other than to mention that significant work exists in the area of CT slice registration and reconstruction but is not necessarily focused on pancreatic imaging [39, 40]. We should also note that registration is necessary to the evaluation of serial (temporal) images of the same patient. For example, in the case of segmentation of the pancreas in multiple, serial scans of a patient that undergoes treatment, registration of CT images obtained at different times may be necessary prior to the assessment of changes from one scan to the next. In the following paragraphs, we will examine each module of the CT image processing algorithm (shown in Fig. 4.5) in more detail.

4.3.1 External Signal Segmentation

Most of the algorithms developed for CT processing involve an initial step of external signal segmentation. This term refers to the removal of signals from the rib cage and spine that usually interfere with the segmentation of internal organs (see Fig. 4.4). These signals have specific characteristics and are usually of higher intensity (pixel value). Their removal is commonly done by thresholding (global and adaptive), edge detection techniques, region growing, and curve fitting [41].

4.3.2 Image Enhancement

This step is usually done to increase the contrast or reduce the noise in an image to allow for more accurate segmentation in the steps that follow. It usually precedes organ segmentation or registration because it offers the potential of redistributing and rescaling pixel values in order to obtain more successful results in the clustering and classification of pixels. Techniques reported in the literature are designed for the spatial or the frequency domain. Spatial domain methods include logarithmic transformations and power law transformations, histogram equalization, image subtraction and averaging, and image smoothing techniques using spatial filters. Frequency (usually Fourier) domain include low-pass filters and high-pass filters for image smoothing and sharpening respectively [41].

4.3.3 Processing—Image Segmentation

Automatic segmentation of CT images admittedly presents significant challenges in computer vision [42]. The primary reason is that the organs are flexible and their size and shape varies as a function of patient characteristics and imaging parameters. Organs are usually accurately localized on CT slices (Fig. 4.4.) but the detection and separation of their boundaries from those of their neighbors and the background is often a difficult task due to the obscure, fuzzy, and irregular edges that are often superimposed by other structures [33, 42]. Even human experts have difficulty in providing unambiguous outlines of the organs' boundaries and consequently present significant inter- and often intraobserver variability, the magnitude of which is a function of experience and training. Historically, standard techniques, such as absolute thresholds, edge detection, and region growing algorithms that perform some type of operation on the gray level distribution of the image pixels are not, by themselves, sufficient for CT segmentation. Combinations of modules, as the one shown in Fig. 4.5, and advanced approaches, e.g., knowledge-based segmentation [42], are necessary to solve this problem.

Several methodologies are reported in the literature for CT slice segmentation although not necessarily focused on pancreatic tissue or pancreatic tumor segmentation. Methods proposed for organ segmentation in CT slices include pixel based (thresholding), edge based, region based, and clustering methods [43]. Interactive segmentation of various organs has also been proposed for 3-D visualization. The reported work used simple thresholding and morphological operations that were interactively controlled by a human user via a 3-D display [44].

There are several free software packages that can be used for the segmentation and registration of CT slices. One of them funded by the National Library of Medicine (NLM) is the Insight Segmentation and Registration Toolkit (ITK) and can be downloaded from www.itk.org. ITK is open-source software that was developed jointly by six principal organizations to support the Visible Human project of NLM. ITK includes several basic segmentation and registration techniques that have been implemented for a variety of medical image analysis applications. In this work, we experimented with several of the methods implemented in ITK. Particularly, region based, threshold select, geodesic active contour segmentation, and fuzzy connectedness with Voronoi classification

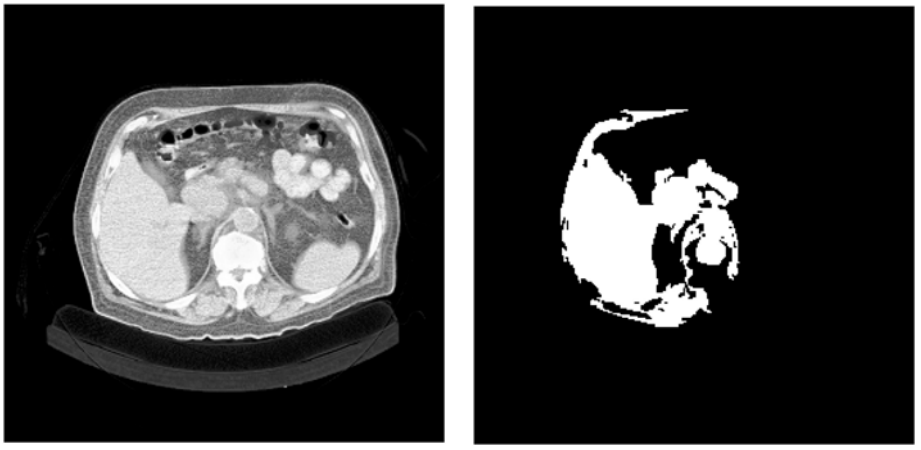


Figure 4.6: (a) Original helical, contrast enhanced CT slice with a tumor at the head of the pancreas indicated by black arrow. (b) Region based segmentation using ITK software on Fig. 4.6(a).

were some of the techniques tested for the segmentation of the pancreas and pancreatic tumors. Initial results suggested that region growing was the best approach because most of the other techniques clustered the majority of the structures in the image together not allowing separation of the pancreas from the other organs. But even with region growing, the pancreas and associated tumor could not be separated from the liver if the pancreatic structures were to remain in the segmented image; separation occurred at the expense of losing most of the information from the gland and associated tumor. Representative segmentation outputs from the region growing approach of ITK are shown in Figs. 4.6(b) and 4.7(b) for two CT slices that contain a mass at the head (Fig. 4.6(a)) and tail (Fig. 4.7(a)) of the pancreas respectively. It should be noted that, although not fully optimized for this application, the tools included in ITK are not likely to yield, by themselves, the desired segmentation outcome because of the low contrast differences between adjacent organs and the way region growing operates. The initial problems we identified in the application of conventional segmentation techniques on CT images of the pancreas include the following:

1. Gray tone segmentation algorithms do not produce accurate regions of the target organ. This is because two different regions of the pancreas or two different organs can have the same or similar gray level tones in CT

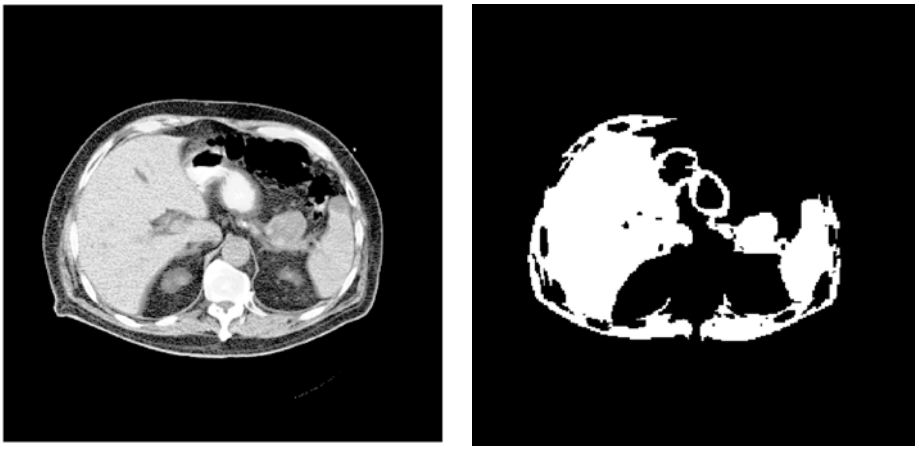


Figure 4.7: (a) Original helical, contrast enhanced CT slice with a pancreatic tumor at the tail of the pancreas indicated by white arrow. (b) Region based segmentation using ITK software on Fig. 4.7(a).

images. Hence, differentiation based on gray level alone is not likely to yield consistent and robust results.

2. The shape of the various organs in the CT slices is not always well defined or consistent from slice to slice. So, it is difficult to select generally applicable characteristics. CAD development is likely to require an adaptive process to deal with this variability.

3. Thresholding techniques based on single global values are not likely to succeed because the gray values of the organs are case-dependent. Gray values depend on the chemical contents of each organ and the physical condition and characteristics of the patient. Gray level normalization may provide a solution to this problem but should be done consistently across slices within the same scan so that it does not prevent registration and reconstruction processes. It should also be done with consideration of the variations among different cases, pathologies, and image sources.

Despite limited performances, however, some of the conventional segmentation techniques, including those implemented in ITK, could be used in the first segmentation step for external signal removal and/or removal of uninteresting structure(s) within the slice, e.g., spleen or kidneys, and isolation of major

organs including the pancreatic areas. This could make the job of subsequent segmentation steps easier and more successful.

4.3.4 Processing—Classification

Very few methodologies have been developed for the classification of pancreatic tumors, e.g., the differentiation between benign and malignant disease or even the differentiation between normal and abnormal pancreas or pancreatic areas reported. One application used several classification schemes to differentiate between pancreatic ductal adenocarcinoma and mass-forming pancreatitis. The methods included artificial neural network classifiers, Bayesian analysis, and Hayashi's quantification method II [45]. The approach used radiologist-extracted CT features for the classification and no automatic segmentation or feature identification was performed. Results indicated that all computer techniques performed similarly to expert radiologists and had no significant benefits [45]. The classification task adds another level of difficult to the segmentation. It is reasonable to hypothesize that classification my be successful if automated feature extraction is performed or when image and nonimage features are merged in the feature set.

4.4 A Novel Algorithm for Pancreatic Tumor Detection and Classification

Fuzzy-based segmentation and classification techniques have been used in various medical imaging applications although not pancreatic cancer [46–48]. An application closest to CT pancreatic imaging with analogous problems is the magnetic resonance imaging (MRI) of the brain and brain tumors. Unsupervised, supervised, and semisupervised fuzzy c-means (FCM) algorithms and knowledge-guided FCM segmentation have been successfully applied to brain tumor MRI applications [49–55]. Similar approaches have also shown promising results for breast tumor segmentation in mammography [56], and lung nodule segmentation on CT images [57]. Here, we present the implementation and initial performance of an FCM based algorithm for pancreatic tumor segmentation and tumor measurements on 2D CT slices. Figure 4.8 presents a flowchart of

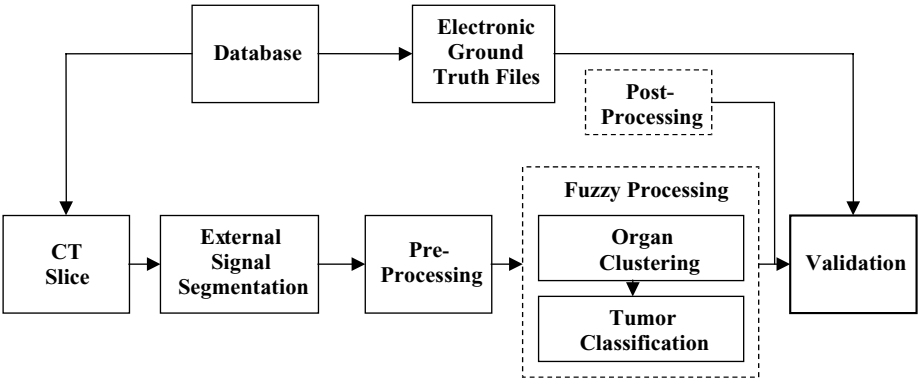


Figure 4.8: Block diagram of CAD algorithm developed for the clustering and classification of pancreatic tumors on helical CT scans.

the design and implementation of the algorithm. The designed CAD scheme follows the general principles presented in Fig. 4.5 but includes additional steps for postprocessing and validation that will be discussed in more detail below.

4.4.1 Medical Image Database

Data collection and database generation is one of the most critical components of algorithm development and one that raises most criticisms and concerns in the scientific community. Datasets are the basis for initial design, training, and testing of the algorithms and together with the analysis tools are key in any validation effort. Given the attention they usually receive and the controversy they sometimes stir, we will review here some of the most important aspects of medical image databases.

Medical imaging data are usually collected retrospectively from completed patient files. Data collection and documentation involves significant amounts of time and effort that is often underestimated. The general guidelines followed in the generation of a database for the particular application of pancreatic cancer are as follows [5, 58–61]:

1. Collect both image and nonimage data and generate complete cases.

2. Review imaging, clinical, and demographic characteristics of the disease and ensure representation of the majority of case types. If this is not

possible, prioritize and focus on selected groups that define the most important clinical problems.

3. Review clinical records of institution to ensure availability and adequacy.

4. Define desired number of cases in overall dataset as well as in subsets needed to address specific problems in addition to the main goal of the effort. Number of cases depends on the training requirements of the selected methodology and the statistical power needs of clinical evaluation studies such as the receiver operating characteristic experiments.

5. Digitize films from analog modalities at the highest possible resolution (spatial and dynamic) and reduce, if necessary, using mathematical interpolation.

6. Generate ground truth information preferably based on pathology information. If this is not possible, use medical experts properly screened to define and outline ground truth on the selected images. Although this approach is subject to high inter- and intraobserver variability, it is often the only possible option. Hence, it is critical for the researcher to develop standardized methods for ground truth file generation, same for all experts, and take any step to eliminate external factors of variability. It is also recommended that all experts' opinions are used in validation instead of the most often occurring response, the union, or overlap of opinions. For computer applications the generation of ground truth information in electronic form is highly desirable and it usually contains outlines of the areas of interest drawn by one or more experts that provides information on the type, location, and size of the area. Electronic ground truth files are discussed in more detail below.

7. Define validation criteria, namely what will be considered as true positive (TP), false positive (FP), true negative (TN), and false negative (FN) for a segmentation or classification outcome. Segmentation validation is usually more demanding and cumbersome process. The existence of specific conventions and consistent criteria in the evaluation of segmentation results is often more important than the variability in ground truth information provided by experts.

8. Collect all imaging information and imaging parameters associated with selected cases.

9. Obtain all available reports, e.g., radiology, pathology, clinical reports, that can assist the researcher in case documentation and evaluation of the database contents.

For our development and preliminary study, data were collected retrospectively from the patient files of the H. Lee Moffitt Cancer Center & Research Institute. Approximately 100 patients undergo a pancreatic CT exam annually at the center. About 2/3 of these patients are diagnosed with pancreatic cancer and about 1/3 with a benign pancreatic mass or cyst. Abdominal scans are also performed for staging patients diagnosed with other cancer types, e.g., breast cancer, that may turn out to be negative for metastatic disease or any disease. Figure 4.9 shows a database design for pancreatic cancer imaging applications.

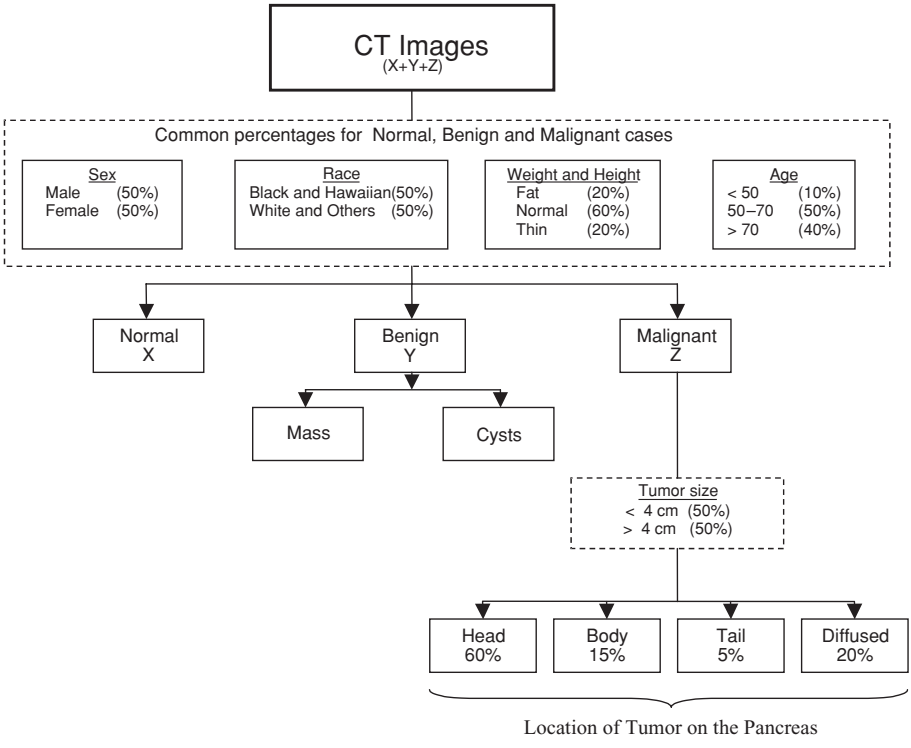


Figure 4.9: Image database design for pancreatic cancer research and CAD development.

The contents of the database, e.g., numbers X, Y, and Z, are determined based on (a) the aims of the project, (b) the clinical characteristics of the pancreatic cancer and benign pancreatic masses, (c) the disease statistics, (d) the demographic characteristics both nationally and locally, (e) the imaging protocols implemented at the Institution, and (f) the requirements of the algorithm design as discussed earlier. Imaging protocols and surveillance procedures may differ among institutions and, hence, CAD goals may differ to accommodate specific clinical practices and requirements. HLMCC's imaging protocol for abdominal helical CT scans of patients diagnosed with or suspected of pancreatic cancer includes three imaging series:

- Series #1: An initial abdominal scan is done with a relatively thick slice (8–10 mm) prior to the administration of contrast material; approximately 5 slices from this series contain information of the pancreas.
- Series #2: An enhanced abdominal scan follows with the same slice thickness as in Series #1 shortly after the intravenous administration of contrast material (a second enhanced scan after a short period of time may also be acquired if requested by the physician). Similar to the first series, approximately 5 slices in this series contain information on the pancreas.
- Series #3: A high-resolution scan of the pancreas at a 4 mm or smaller slice thickness. This scan is not routinely performed and depends on the patient and the physician. This series consists of about 10 slices through the pancreas.
- Series #4: A renal delay scan that acquires images through the kidneys only. This series includes partial information on the pancreas.

Series #1 and #4 are not likely to be of value at least in the initial algorithm development because pancreatic tumors are clinically evaluated in contrast enhanced scans, i.e., Series #2, and insufficient information is present in Series #4.

In addition to the CT images and imaging parameters, the following information was also collected or generated: (a) radiology reports, (b) pathology reports, (c) demographic information, (d) other nonimage information including lab tests, and (e) electronic ground truth files. All data were entered in a relational database that links image and nonimage information. All patient

identifiers were removed prior to any research and processing to meet confidentiality requirements.

4.4.2 Electronic Ground Truth File Generation

Ground truth files are and should be generated by physicians that are experts in the specific imaging modality and anatomy, e.g., CT interpretation and pancreatic disease. Despite subjectiveness and variability concerns, manual outlines by human experts are often the only way to establish ground truth or a form of ground truth to which image segmentation techniques can be compared to. If pathology information is available, as in the case of patients undergoing surgery or biopsy, then it can be used to provide stronger evidence on the true tumor size and possibly shape although the two sources of information, radiology and pathology, are not directly correlated. Unfortunately, in the case of pancreatic cancer, many of the tumors will not be resectable and will only undergo treatment; actually these are the patients one is most interested in, if one wants to determine treatment effects and tumor response over time. As a result, manual outlines by experts are the only option for ground truth generation on helical CT scans of the pancreas.

Figures 4.10(a) and 4.10(b) show two tumor outlines generated independently for the same CT slice and the same case with a malignant mass at the

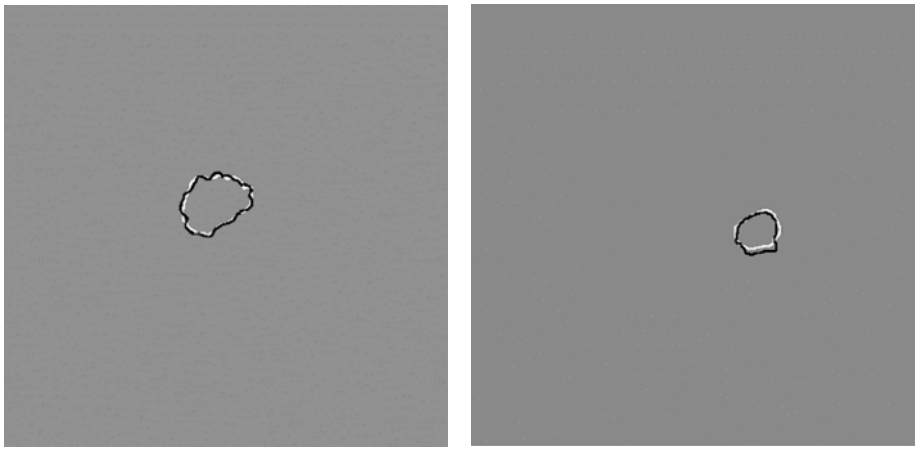


Figure 4.10: (a) Two manual outlines of the pancreatic tumor shown in Fig. 4.6(a). (b) Two manual outlines of the pancreatic tumor shown in Fig. 4.7(a).

head and tail of the pancreas respectively (original images shown in Figs. 4.6(a) and 4.7(a)). These outlines are part of the ground truth files generated for the CT slices and used for segmentation validation. Variations in the outlines as the ones seen in Fig. 4.10 are expected and inevitable between experts and could make segmentation validation a strenuous task. Often there is no right or wrong answer and it is our recommendation that both are considered in an evaluation process.

Measures can and should be taken to increase the accuracy of this information and at a minimum remove external sources of variability or error. These measures include the following:

(i) Establish optimum and standard viewing and outline conditions in terms of monitor display and calibration, ambient light, manual segmentation tool(s), and image manipulation options.

(ii) Use all individual manual outlines for evaluation as a possible way to account for expert variability. For example, use both outlines shown in Fig. 4.10 for segmentation validation. Alternative options are to determine the union or overlap of outlines or use a panel of experts to obtain a consensus on one outline per image.

(iii) Provide all available information to the expert before he/she generates truth file.

(iv) Have experts perform initial outlines independently to avoid bias (any joint outlines are done in addition to the originals).

(v) Review the expertise of the "experts" and their physical condition prior to the initiation of the process (number of cases read within a certain time frame, familiarity with computer tools, training, fatigue).

(vi) Establish standard criteria and conventions to be followed by all experts in their outlines.

Ground truth files are generated for all cases in the designed database but for a selected number of image series and slices to reduce physician effort. Specifically, in the cases where there is no high-resolution series (#3), ground truth files are generated for the 5 enhanced CT slices of Series #2 that contain the pancreas. In the cases where the high-resolution series (#3) is available,

ground truth files are generated for both the 8 mm slices of the pancreas in Series #2 and every other slice in Series #3 (10 slices total); "ground truth" for the intermediate slices of Series #3 is obtained by interpolation. (Ground truth could be extrapolated from the 4 mm slices to the 8 mm slices but slice registration would be required prior to this process.)

Truth files are images of the same size as the original slice and include (a) the location of the pancreas and outline of its shape; (b) the location of the pancreatic tumor(s), masses, or cysts, and their shape outline(s) (Fig. 4.10); (c) the location and identification of neighboring organs and their shape outlines; (d) the identification of any vascular invasion and metastases sites and outlines. Truth files are generated using a computer mouse to outline the areas of interest on CT slices that are displayed on a high resolution (2048×2560 pixels) and high luminance computer monitor. Pixels in the ground truth files are assigned a specific value that corresponds to an outlined organ or structure, e.g., a gray value of 255 is assigned to the pixels that correspond to the outline of the pancreatic tumor(s), a gray value of 200 to the pixels that correspond to the outline of the normal pancreas, etc.

4.4.3 External Signal Segmentation

The approach that was implemented was based on edge detection, line tracing, and histogram thresholding techniques [43]. The requirements for this process do not differ significantly from those followed in standard chest radiography (CXR) and several of the concepts described in CXR literature are applicable to CT as well [62]. One primary issue in this module was the desired level of accuracy in the removal of the external signals, i.e., signals from the rib cage and spine. Increasing the accuracy level, increased the computational requirements and the complexity of the methodology. Figures 4.11 and 4.12 show the external signal removal for the slices of Figs. 4.6(a) and 4.7(a).

A histogram equalization approach was used to remove the regions that correspond to the rib cage and spine that usually are the highest intensity regions in the image. Points on the rib cage were defined using the pixel characteristics of the rib cage and these points were interpolated using a spline interpolation technique [63]. The boundary of the rib cage was then estimated and removed.

Figure 4.11: External signal segmentation for the slice of Fig. 4.6(a).

4.4.4 Preprocessing—Enhancement

Our enhancement approach aimed at increasing the image contrast between the pancreas and organs in close proximity. A histogram equalization approach was implemented for this purpose and yielded satisfactory results (Gaussian and Wiener filters seemed to benefit these images as well) [64]. Wavelet-based enhancement was also considered as an alternative option for removing unwanted background information and better isolating the signals of interest [65]. The method was promising but may present an issue when used in combination with registration or reconstruction processes.

Enhancement generally benefits CAD algorithms but in 3-D imaging modalities like CT, it may have an adverse effect on the registration of the 2-D data, if it is not uniformly done across slices. Wavelet-based enhancement may worsen

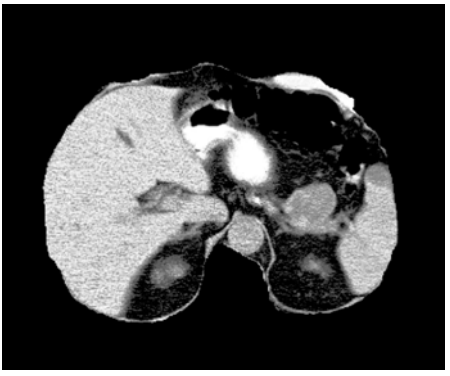


Figure 4.12: External signal segmentation for the slice of Fig. 4.7(a).

the situation since it operates in the frequency domain and may not necessarily preserve the spatial features of the CT images as needed for registration. A standardization or normalization method may offer a solution in regaining all spatial information when transforming from the frequency back to the spatial domain. However, no such method was established for this application or is readily available. If registration is not part of the process, the enhancement step could significantly benefit subsequent clustering and classification on the CT images [66].

4.4.5 Fuzzy Clustering

A cluster or a class may be defined as a set of objects with similar characteristics (features) and different from other objects outside the group. When data are clustered, a condition is chosen for measuring the similarity of an object to that cluster. There are three types of clustering techniques: crisp, also known as hard or classical clustering, fuzzy, and probabilistic clustering [43, 48].

Crisp or classical clustering algorithms classify objects as part of a cluster or not part of a cluster. The object usually takes a hard value of either 0 or 1 with respect to a cluster or class. Cluster separation is very clear and no overlap is allowed. In real applications, however, and medical imaging in particular, there is overlap between the various data types and there are no crisp boundaries between classes.

Fuzzy clustering algorithms allow for data overlap because they use class membership values instead of binary assignments. These algorithms treats an object, e.g., an image pixel, as part of all clusters. The object is given a weight with respect to each cluster and the weight vector is unconstrained taking a value between 0 and 1. A weight greater than 0.5 indicates that the object is more likely to be similar to that cluster.

Probabilistic clustering is similar to fuzzy clustering with the exception of the weight vector being constrained in this case. Namely, the sum of all weights assigned to an object equals 1.

Various fuzzy clustering models are proposed in the literature. One widely used model is the fuzzy c-means (FCM) algorithm developed by Bezdek [48]. We have implemented FCM in several variations mainly for MRI brain tumor classification. Variations include unsupervised FCM [67, 68], FCM combined with knowledge based clustering [51], FCM combined with validity guided

(re)clustering [50], semisupervised FCM (ssFCM) algorithms [49], and supervised FCM algorithms [52, 53]. In supervised learning, FCM is trained under complete supervision, namely the algorithm is forced to learn every class correctly. In unsupervised learning, data are clustered according to their similarity, there is no forced training, and the algorithm is allowed to make its own classification decision. Semisupervised learning offers a middle path between the previous two. In this case, the user defines a degree of supervision that leaves some room for the algorithm to make its own decisions while not entirely unrestricted to do so. Semisupervised learning may offer advantages for the clustering of data with significant overlap. In the following paragraphs we will discuss few key theoretical aspects of the three learning approaches.

The FCM family of objective functions is defined as [47]

$$J(U, V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m \|x_k - v_i\|_A^2$$

where $m \in [1, \infty)$ is a weighing exponent on each fuzzy membership, $U \in M_{\text{fcn}}$ is a constrained fuzzy c-partition of the dataset X of n feature vectors x_j in c clusters, $V = (v_1, v_2, \ldots, v_c)$ are c-vector prototypes in the p-dimensional feature space R^p and A is any positive definite $(p \times p)$ matrix. U and V may minimize J only if

$$u_{ik} = \left[\sum_{j=1}^{c} \left(\frac{\|x_k - v_i\|_A}{\|x_k - v_j\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \quad 1 \le i \le c \quad 1 \le k \le n \tag{4.1}$$

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \quad 1 \le i \le c \tag{4.2}$$

Note that U is a $(c \times n)$ matrix of the u_{ik} values and $\|x_k - v_i\|_A^2 = (x_k - v_i)^T A (x_k - v_i)$. The steps followed in the implementation of the FCM algorithm are as follows [47]:

1. Input unlabeled dataset $X = \{x_1, x_2, \ldots, x_n\}$
2. Choose parameters c (number of clusters), T, A, $m \ge 1$, and $\varepsilon > 0$
3. Initialize $U_0 \in M_{fcn}$ randomly
4. Compute the $\{v_{i,0}\}$ from (2) for $1 \le i \le c$
5. For $t = 1, 2, \ldots, T$ do

(i) computer $\{u_{ik,t}\}$ from (1) for $1 \leq K \leq n$

(ii) computer error as $\|U_t - U_{t-1}\|$

(iii) if error $\leq \varepsilon$ stop; else compute $\{v_{i,t}\}$ from (2)

(iv) continue to the next t

In an unsupervised method, all n label vectors are unknown during initialization. Once a condition is satisfied and the parameter c is selected, the algorithm stops and each row in the matrix U corresponds to a cluster that may or may not have a label [47]. To avoid problems associated with nonlabeled clusters, a supervised FCM may be used. To change an unsupervised to a supervised method, expert human intervention is required and two things may be done:

1. The operator selects the number of clusters c during initialization.
2. The operator assigns labels to the classes at termination.

The problematic aspects of a supervised approach are the need for high-quality labeled data and the variability introduced by human intervention. An alternative to the supervised and unsupervised versions of FCM is the semisupervised method where a small set of the data is labeled but the majority of the data is unlabeled. In ssFCM, we let

$$X = \left\{ \underbrace{x_1^1, x_2^1, \ldots, x_{n_1}^1, x_1^2, x_2^2, \ldots, x_{n_2}^2 \ldots, x_1^c, x_2^c, \ldots, x_{n_c}^c}_{\text{labeled data}} \middle\| \underbrace{x_1^u, x_2^u, \ldots, x_{n_u}^u}_{\text{unlabeled data}} \right\}$$

denote partially labeled data; superscripts denote the class label and n_i is the number of samples having the same label vector in the partition matrix U. Using the labeled set of data we find the center of the clusters iteratively until the terminating condition is satisfied. The unlabeled data are then introduced for finding the cluster centers. This method is more stable as the centers are well defined from the labeled data used for training. The clusters are also well defined with the correct physical data labels.

The cluster centers for the labeled data are calculated as

$$v_{i,0} = \frac{\sum_{k=1}^{n_d} (u_{ik,0}^d)^m x_k^d}{\sum_{k=1}^{n_d} (u_{ik,0}^d)^m} \quad 1 \leq i \leq c \tag{4.3}$$

where the subscript d denotes design or training data. Having the labels for the

n_{d} points, we need to update only the n_{u} columns in U by calculating:

$$u^{u}_{ik,t} = \left[\sum_{j=1}^{c} \left(\frac{\|x^{u}_{k} - v_{i,t-1}\|_A}{\|x^{u}_{k} - v_{j,t-1}\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \quad 1 \le i \le c \;\; 1 \le k \le n_{\mathrm{u}} \;\; t = 1, 2, \ldots, T \tag{4.4}$$

Once the initial cluster centers $\{v_{i,0}\}$ are calculated, cluster centers are recomputed using the unlabeled data as

$$v_{i,t} = \left(\frac{\sum_{k=1}^{n_{\mathrm{d}}} (u^{d}_{ik,t})^m x^{d}_{k} + \sum_{k=1}^{n_{\mathrm{u}}} (u^{u}_{ik,t})^m x^{u}_{k}}{\sum_{k=1}^{n_{\mathrm{d}}} (u^{d}_{ik,t})^m + \sum_{k=1}^{n_{\mathrm{u}}} (u^{u}_{ik,t})^m} \right) \quad 1 \le i \le c \;\; t = 1, 2, \ldots, T \tag{4.5}$$

where the subscript u is now used to denote the unlabeled data contribution.

In practice, n_{d} is much smaller than n_{u}. For example, an abdominal CT image is usually 512×512 or 262,144 pixels. A pancreatic mass of about 4 cm in maximum diameter may cover approximately 1200 pixels in an image with a resolution of 1 mm/pixel while the pancreas itself may be up to 5000 pixels. Neighboring major organs and structures may be up to 20,000 pixels depending on the slice. A very small percentage of these pixels will be labeled in the ssFCM approach. For example for $c = 4$ a quarter of the pixels in each class per slice are labeled, which approximately equals to $n_{\mathrm{d}} = 8000$ and $n_{\mathrm{u}} = 254{,}144$.

To reduce potential bias that may be introduced by large differences between n_{d} and n_{u} as well as between the difference tissue classes, one more modification of c-means is introduced that allows us to weigh the fewer labeled samples more heavily than their unlabeled counterparts. Furthermore, such weighing allows us to assign much larger weights to small clusters to better separate them from larger ones. This is done by introducing weights $W = (w_1, w_2, \ldots, w_{nd})$ in the Eq. (4.5) as

$$V_{i,t} = \left(\frac{\sum_{k=1}^{n_{\mathrm{d}}} w_k (u^{d}_{ik,t})^m x^{d}_{k} + \sum_{k=1}^{n_{\mathrm{u}}} (u^{u}_{ik,t})^m x^{u}_{k}}{\sum_{k=1}^{n_{\mathrm{d}}} w_k (u^{d}_{ik,t})^m + \sum_{k=1}^{n_{\mathrm{u}}} (u^{u}_{ik,t})^m} \right) \quad 1 \le i \le c \;\; t = 1, 2, \ldots, T \tag{4.6}$$

In general, W is a vector of positive real numbers and when $w_{nd} = 1$ then ssFCM becomes standard unsupervised FCM [47].

The implementation of an ssFCM algorithm involves the following steps:

1. Given partially labeled data $X = X^d \cup X^u$, set the n_{d} and n_{u} parameters as $n_{\mathrm{d}} = |X^d|$ and $n_{\mathrm{u}} = |X^u|$; the c parameter is known and fixed by the training data.

2. Choose W, T, A, $m > 1$, and $\varepsilon > 0$.
3. Initialize $U_0 = [U^d | U_0^u]$, with $U_0^u \in M_{\text{fcn}}$.
4. Compute initial cluster centers using Eq. (4.3).
5. For $t = 1, 2, \ldots, T$ do the following:
   (i) compute the $u_{ik,t}^u$ from Eq. (3.4)
   (ii) compute error as $\|U_t^u - U_{t-1}^u\|$
   (iii) if error is $\leq \varepsilon$ stop; else compute $v_{i,t}$ from Eq. (4.6)
   (iv) continue to the next t

In ssFCM, centers of smaller clusters may be prevented from migrating toward larger clusters by giving high weights to the training data that correspond to the smaller clusters.

All FCM versions described previously have been implemented in our laboratory. Preliminary results for pancreatic cancer imaging have been obtained with the unsupervised FCM algorithm and these will be discussed in the following sections. For our pilot work, the number of clusters was set to 4 and FCM was applied in two stages. Specifically, in the first run, FCM was applied to the CT image after the external signals were removed to cluster all major organs in the same class as shown in Figs. 4.13 and 4.14 for the original slices of Figs. 4.6(a) and 4.7(a). In these figures, all major organs, including the pancreas, are grouped in the same class labeled as 1. Class 1 pixels were then remapped to the pixel values in the original CT image and FCM was applied to this selected

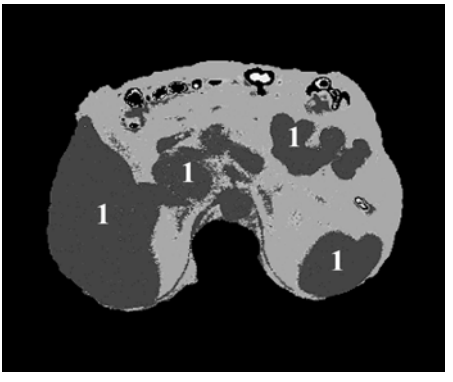


Figure 4.13: Unsupervised FCM on Fig. 4.6(a).

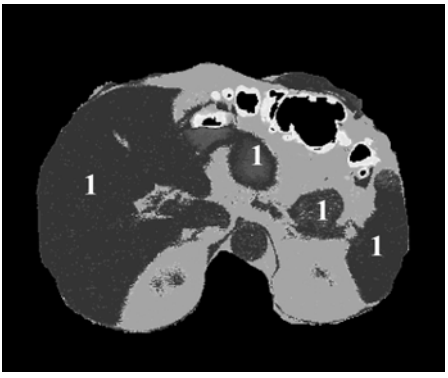


Figure 4.14: Unsupervised FCM on Fig. 4.7(a).

area for a second time using the same parameters. The output of the second run is shown in Figs. 4.15 and 4.16. Although not immediately evident in a grayscale representation, four pixel clusters were identified in these images. One of these clusters (encompassed by a white outline) corresponds to the pancreatic tumors indicated by arrows in Figs. 4.6(a) and 4.7(a) and the pancreas.

Several issues remain to be addressed in this FCM application including (a) what is the appropriate number of clusters to differentiate between tumor and nontumor pancreatic areas, (b) how many labeled data are needed for

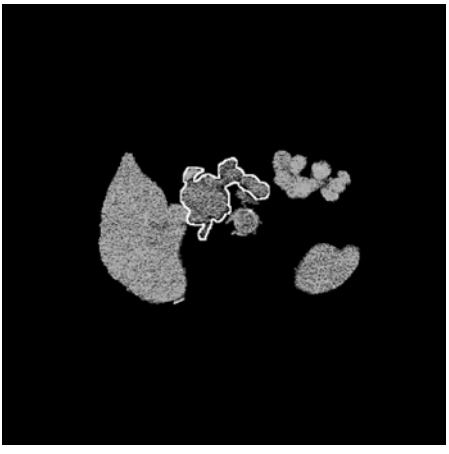


Figure 4.15: Segmentation results obtained from applying the FCM algorithm to the area defined by class 1 pixels in Fig. 4.13. The white outline indicates the cluster that corresponds to the pancreas and tumor pixels.

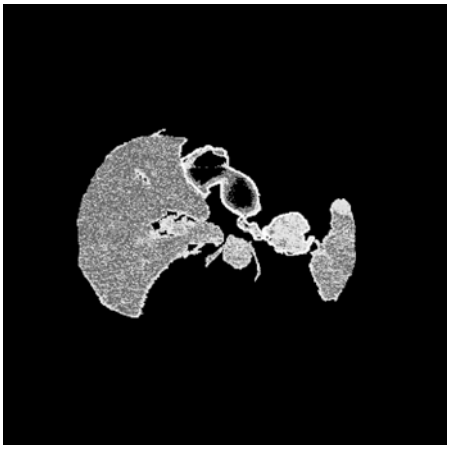


Figure 4.16: Segmentation results obtained from applying the FCM algorithm to the area defined by class 1 pixels in Fig. 4.14. The white outline indicates the cluster that corresponds to the pancreas and tumor pixels.

training to reduce misclassification, (c) in what proportion the weights should be assigned for each cluster, (d) what is the optimum stopping criteria, ϵ, (e) can we avoid the two-stage application by removing unwanted signals further or by reducing the region of interest on which to apply FCM, and (f) validation of the clustering results relative to ground truth files.

Finally, a semisupervised algorithm is currently applied for pancreatic tumor clustering. In this ssFCM, a small subset of feature vectors (or pixels) is labeled by a radiologist, expert in CT interpretation, and used as training information to define the clusters. The disadvantage of ssFCM is that it requires the expert's input through the selection of training pixels for the tissue classes in an image. On the other hand, ssFCM is less likely to be adversely affected by imperfect training data than its fully unsupervised counterpart and it may be more efficient, i.e., it could achieve better results than FCM for the same number of classes (4). From a practical perspective, and assuming that the results from this algorithm turn out to be more robust and consistent, it is not clinically impractical and may also be advantageous to have the expert initiate the ssFCM process during CT image review. Alternative methods under consideration is the optimization of FCM by combining it with a validity-guided clustering technique [50] or a knowledge-based system [54], or a genetic algorithm technique [55] that have shown potential for improvement. Finally, the imaging characteristics of the

pancreatic tumors and nonimage, demographic information could be merged to guide cluster initialization and tissue classification. All these options are topics of future investigations.

4.4.6 Validation

Our pilot study on pancreatic cancer did not include a validation step due to the small number of the tested images to date. However, the evaluation of the clustering and segmentation outputs is expected to be a major part of this application. Hence, we will close our algorithm description with few remarks on segmentation validation issues and a summary of the measures proposed for this purpose.

Validation requires a gold standard segmentation that represents the "absolute truth" on the size and shape of the object of interest. The lack of a gold standard or absolute ground truth in most medical imaging applications does not allow an absolute quantitative evaluation of the segmentation output. The best and often only, option available is segmentations generated by expert observers that may be biased and also exhibit significant inter- and intraobserver variability. In some cases, an alternative approach to the direct evaluation of segmentation results is the use of simulation or phantom studies, [69] the use of relative performance measures, or the use of classification outcomes [70].

The goal of validation in our application is to demonstrate that the automatic methods proposed for the segmentation of pancreatic tumors will lead to standardized and more reproducible tumor measurements than the manual and visual estimates performed traditionally by experts. Tumor size, area, and volume are parameters currently used to determine tumor resectability, and response to treatment. Greater accuracy, less variability, and greater reproducibility in these measurements is expected to have a significant impact on the diagnosis and treatment of pancreatic cancer [71].

An indicated in Fig. 4.8, a postprocessing step is usually applied to the clustered data prior to validation in order to generate smooth contours of the organs and tumors that can then be compared to those in the truth files; see for example the truth files in Fig. 4.10 and the FCM segmentations (white outlines) of Figs. 4.15 and 4.16). From the measures available for segmentation validation, [72] we have selected and implemented those that are recommended for medical imaging applications and are particularly suited for the comparison of

computer-generated to hand-drawn boundaries [73–75]. In addition, they are relatively computationally efficient and are not limited to specific shape patterns. These measures are as follows [74]:

1. The Hausdorff distance $h(A, B)$ between two contours of the same object (tumor), one generated by an expert (A) and one generated by the computer (B).

   Let $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ be the set of points on the two contours (each point representing a pair of x and y coordinates) then the distance of a point a_i to the closest point on curve B is defined as

   $$d(a_i, B) = \min_j \|b_j - a_i\|$$

   Similarly the distance of a point b_j to the closest point on curve A is defined as

   $$d(b_j, A) = \min_i \|a_i - b_j\|$$

   The Hausdorff distance $h(A, B)$ is defined as the maximum of the above distances between the two contours, i.e.

   $$h(A, B) = \max\left(\max_i\{d(a_i, B)\}, \max_j\{d(b_j, A)\}\right).$$

2. The degree of overlap OL between the areas G and E encompassed by contours A and B. The overlap is defined as the ratio of the intersection and the union of the two areas, i.e, the ground truth area G and the experimental computer generated area E:

   $$\mathrm{OL} = \frac{G \cap E}{G \cup E}$$

   The ratio is 1 if there is perfect agreement and 0 if there is complete disagreement.

3. The mean absolute contour distance (MACD). MACD is a measure of the difference between the two contours. To estimate MACD, a one-to-one correspondence between the points of the two curves is required. Once this correspondence is established, the distances between corresponding points are estimated; their average corresponds to MACD. In addition to the absolute differences entering the MACD calculation, the signed distances between the curves may also be computed and used to determine the bias

of an algorithm or any regional effects on the segmentation process, i.e., pancreatic areas closer to the liver may be less accurately segmented than areas away from large organs [74].

The first two measures above are sensitive to the size and shape of the segmented objects and also depend on the image spatial resolution. The third measure is independent of object size and image resolution and preferred if images from different sources are to be compared.

Alternatively to custom-made routines, the VALMET segmentation validation software tool that is publicly available could be used to generate these metrics in 2D and 3D [73]. Tools such as VALMET and ITK may offer the standardization missing from the validation of segmentation algorithms and reduce variability. Currently, there is no agreement on the "best method" or "best methods" for analyzing and validating segmentation results. The need for standardized measures that are widely acceptable is significant as is the need for establishing conventions on how to use expert-generated ground truth data in the evaluation process.

In a final note, the reader is reminded that a statistical analysis that measures the agreement between the measured parameters from different segmentation algorithms or the agreement between computer and observer performances should be part of the validation process. Computer and expert data are compared with a variety of statistical tools. The most frequently reported ones include (a) linear regression analysis to study the relationship of the means in the various segmentation sets [76, 77], (b) paired t test to determine agreement between the computer method(s) and the experts [76, 77], (c) Williams index to measure interobserver or interalgorithm variability in the generation of manual outlines [74], and (d) receiver operating characteristic analysis and related methods to obtain sensitivity and specificity indices by estimating the true positive and false positive fractions detected by the algorithm and/or the observer [78].

4.5 Conclusions

This chapter discussed aspects related to the segmentation of medical images for the purpose of tumor evaluation and treatment assessment. Pancreatic cancer imaging by CT was used as the basis for discussing image segmentation

issues for medical imaging and CAD applications. It was also used in an effort to open the pancreatic cancer imaging area into possibly more research and discussions considering that it is relatively under-investigated and unknown despite its significant toll on health care.

The current state-of-the-art in CAD methodologies for CT and pancreatic cancer was reviewed and limitations were discussed that led to the development of a novel, fuzzy logic-based algorithm for the clustering and classification of pancreatic tumors on helical CT scans. This algorithm was presented here and its pilot application on selected CT images of patients with pancreatic tumors was used as the basis to discuss issues associated with tumor segmentation and validation of the results.

The problems and difficulties encountered today by the radiologists and the oncologists dealing with pancreatic carcinoma are numerous and they are often associated with the limitations of the current imaging modalities, the observer biases, and the inter- and intraobserver variability. Among the most striking weaknesses is the inability to detect small tumors, to consistently differentiate between pancreatic tumors and benign conditions of the pancreas putting the patient through several imaging procedures and medical tests, to accurately measure tumor size and treatment effects.

Computer tools could play a diverse role in pancreatic cancer imaging. The primary goal of the system presented here was the automated segmentation of the normal and abnormal pancreas and associated pancreatic tumors from CT images. However, these tools could have a broader and more diverse role in the detection, diagnosis, and management of this disease that could change the current standard of care. Among other applications, CAD methodologies could provide objective measures of pancreatic tumor size and response to therapy that will allow (a) accurate and timely assessment of tumor resectability, (b) accurate and timely estimates of tumor size as a function of time and treatment, and (c) standardized evaluation and interpretation of tumor size and response to treatment. CAD techniques could further lead to 3-D reconstructions of the pancreas and tumors and impact surgery and radiation treatment.

Validation is and should be a major part of CAD development and implementation. Medical imaging applications, however, present unique problems to CAD validation, e.g., lack of an absolute gold standard, lack of standardized statistical analysis and evaluation criteria, time-consuming and costly database generation procedures, and other. Yet, CAD researchers are asked to find ways to overcome

limitations and properly validate medical CAD algorithms including those that involve segmentation or clustering. Several options have been proposed in this chapter for this purpose. As we learn more about this area, however, we find that it may be possible to define a new family of validation criteria better suited for medical imaging applications. These criteria are likely to link algorithm performance to actual clinical outcomes. We could use, for example, classification results as a measure of segmentation performance.

4.6 Acknowledgments

We acknowledge the valuable contributions of Dr. Amine Bensaid in the implementation of the fuzzy algorithms for this application. We are also thankful to Angela Salem, Joseph Murphy, Isaac E. Brodsky, and Deepakchandran Chinnaswami for their indispensable assistance in data collection and image processing, as well as in the preparation of this manuscript.

Questions

1. *What are the physical characteristics of helical CT scans that may impact CAD algorithm design and performance?*
2. *What are the clinical characteristics of pancreatic cancer that may impact CAD algorithm design?*
3. *What is the general approach for image segmentation of medical images?*
4. *Advantages and disadvantages of unsupervised, supervised, and semi-supervised clustering methodologies for image segmentation.*
5. *What is FCM and when is it used for image segmentation? List any advantages over classical segmentation techniques.*
6. *What are the differences between FCM and ssFCM? List advantages and disadvantages of the two techniques.*
7. *List methods that can be used for the optimization of FCM for image segmentation.*

8. *What are the metrics used for the validation of a segmentation output?*
9. *What are the major limitations and problems associated with the validation of segmentation algorithms in medical imaging applications?*
10. *What are the statistical tools used for the analysis of the segmentation results including tools to determine the agreement between different algorithms and observers or within groups.*

Bibliography

[1] Jemal, A., Thomas, A., and Murray, T., Cancer statistics, 2002, CA Cancer J. Clin., Vol. 52, pp. 23–47, 2002.

[2] Kern, S., Tempero, M., and Conley, B., (Co-Chairs), Pancreatic cancer: An agenda for action, Report of the Pancreatic Cancer Progress Group, National Cancer Institute, February 2001.

[3] Kuvshinoff, B. W. and Bryer, M. P., Treatment of resectable and locally advanced pancreatic cancer, Cancer Control, Vol. 7, No. 5, pp. 428–436, 2000.

[4] Lin, Y., Tamakoshi, A., Kawamura, T., Inaba, Y., Kikuchi, S., Motohashi, Y., Kurosawa, M., and Ohno, Y., An epidemiological overview of environmental and genetic risk factors of pancreatic cancer, Asian Pacific J. Cancer Prev., Vol. 2, pp. 271–280, 2001.

[5] Li, D. and Jiao, L., Molecular epidemiology of pancreatic cancer, Int. J. Gastrointest. Cancer, Vol. 33, No. 1, pp. 3–14, 2003.

[6] Ghadirian, P., Lynch, H. T., and Krewski, D., Epidemiology of pancreatic cancer: an overview, Cancer Detect Prev., Vol. 27, No. 2, pp. 87–93, 2003.

[7] Van Hoe, L. and Baert, A. L., Pancreatic carcinoma: Applications of helical computed tomography, Endoscopy, Vol. 29, pp. 539–560, 1997.

[8] Yeo, T. P., Hruban, R. H., Leach, S. D., Wilentz, R. E., Sohn, T. A., Kern, D. E., Iacobuzio-Donahue, C. A., Maitra, A., Goggins, M., Canto, M. I., Abrams, R. A., Laheru, D., Jaffee, E. M., Hidalgo, M., and Yeo, C. J., Pancreatic cancer, Curr. Prob. Cancer, Vol. 26, No. 4, pp. 176–275, 2002.

[9] Tamm, E. P., Silverman, P. M., Charnsangavej, C., and Evans, D. B., Diagnosis, staging, and surveillance of pancreatic cancer, AJR, Vol. 180, pp. 1311–1323, 2003.

[10] Clark, L. R., Jaffe, M. H., Choyke, P. L., Grant, E. G., and Zeman, R. K., Pancreatic imaging, Radiol. Clin. North Am., Vol. 23, No. 3, pp. 489–501, 1985.

[11] Haaga, J. R., Alfide, R. J., Zelch, M. G., Meany, T. F., Boller, M., Gonzalez, L., and Jelden, G. L., Computed tomography of the pancreas, Radiology, Vol. 120, pp. 589–595, 1976.

[12] Haaga, J. R., Alfide, R. J., Harvilla, T. R., Tubbs, R., Gonzalez, L., Meany, T. F., and Corsi, M. A., Definitive role of CT scanning of the pancreas: The second year's experience, Radiology, Vol. 124, pp. 723–730, 1977.

[13] Sheth, S., Hruban, R. K., and Fishman, E. K., Helical CT of islet cell tumors of the pancreas: Typical and atypical manifestations, AJR, Vol. 179, pp. 725–730, 2002.

[14] Horton, K. M. and Fishman, E. K., Adenocarcinoma of the pancreas: CT imaging, Radiol. Clin. North Am., Vol. 40, pp. 1263–1272, 2002.

[15] Horton, K. M., Multidetector CT and three-dimensional imaging of the pancreas: state of the art, J. Gastrointest. Surg., Vol. 6, pp. 126–128, 2002.

[16] Winston, C. B., Mitchell, D. G., Outwater, E. K., and Ehrlich, S. M., Pancreatic signal intensity on T1-weighted fat saturation MR images: Clinical correlation, J. Magn. Reson. Imaging, Vol. 5, pp. 267–271, 1995.

[17] Ragozzino, A. and Scaglione, M., Pancreatic head mass: What can be done? Diagnosis: Magnetic resonance imaging, J. Pancreas, Vol. 1, pp. 100–107, 2000.

[18] Barish, M. A., Yucel, E. K., and Ferrucci, J. T., Magnetic resonance cholangiopancreatography, NEJM, Vol. 341, pp. 258–264, 1999.

[19] Fulcher, A. S. and Turner, M. A., MR pancreatography: A useful tool for evaluating pancreatic disorders, Radiographics, Vol. 19, pp. 5–24, 1999.

[20] Adamek, H. E., Albert, J., Breer, H., Weitz, M., Schilling, D., and Riemann, J. F., Pancreatic cancer detection with magnetic resonance cholangiopancreatography and endoscopic retrograde cholangiopancreatography: a prospective controlled study, Lancet, Vol. 356, pp. 190–193, 2000.

[21] Mertz, H. R., Sechopoulos, P., Delbeke, D., and Leach, S. D., EUS, PET, and CT scanning for evaluation of pancreatic adenocarcinoma, Gastrointest. Endosc., Vol. 52, pp. 367–371, 2000.

[22] Wiersema, M. J., Accuracy of endoscopic ultrasound in diagnosing and staging pancreatic carcinoma, Pancreatology, Vol. 1, pp. 625–632, 2001.

[23] Kalra, M. K., Maher, M. M., Boland, G. W., Saini, S., and Fischman, A. J., Correlation of positron emission tomography and CT in evaluating pancreatic tumors: Technical and clinical implications, AJR, Vol. 181, No. 2, pp. 387–393, 2003.

[24] Koyama, K., Okamura, T., Kawabe, J., Nakata, B., Hirakawa-Chung, K. Y. S., Ochi, H., and Yamada, R., Diagnostic usefulness of FDG PET for pancreatic mass lesions, Ann. Nuclear Med., Vol. 15, No. 3, pp. 217–224, 2001.

[25] Dupuy, D. E., Costello, P., and Ecker, C. P., Spiral CT of the pancreas, Radiology, Vol. 183, pp. 815–818, 1992.

[26] DiChiro, G. and Brooks, R. A., The 1979 Nobel prize in physiology and medicine, Science, Vol. 206, No. 30, pp. 1060–1062, 1979.

[27] Kalender, W. A. and Polacin, A., Physical performance characteristics of spiral CT scanning, Med. Phys., Vol. 18, No. 5, pp. 910–915, 1991.

[28] Boone, J. M., Computed tomography: Technology update on multiple detector array scanners and PACS considerations, In: Practical Digital Imaging and PACS, Seibert, J. A., Filipow L. J., and Andriole, K. P., eds., AAPM Medical Physics Monograph No. 25, Medical Physics Publishing, Madison, WI, pp. 37–65, 1999.

[29] Swindell, W. and Webb, S., X-ray transmission computed tomography, In: The Physics of Medical Imaging, Webb, S., ed., Adam Hilger, Bristol, pp. 98–127, 1988.

[30] McCollough, C. H. and Zink, F. E., Performance evaluation of a multislice CT system, Med. Phys., Vol. 26, No. 11, pp. 2223–2230, 1999.

[31] Sheedy, P. F., II., Stephens, D. H., Hattery, R. R., MacCarty, R. L., and Williamson, B., Jr., Computer tomography of the pancreas, Radiol. Clin. North Am., Vol. 15, No. 3, pp. 349–366, 1977.

[32] Dendy, P. P. and Heaton, B., Physics for Diagnostic Radiolog, 2nd ed., Medical Science Series, Institute of Physics Publishing, Bristol, 1999.

[33] Remer, E. M. and Baker, M. E., Imaging of chronic pancreatitis, Radiol. Clin. North Am., Vol. 40, pp. 1229–1242, 2002.

[34] Love, L., (guest ed.), Symposium on abdominal imaging, Radiol. Clin. North Am., Vol. 17, No. 1, 1979.

[35] Frank Miller, H., (guest ed.), Radiology of the pancreas, gallbladder, and biliary tract, Radiol. Clin. North Am., Vol. 40, No. 6, 2002.

[36] Sheth, S. and Fishman, E. K., Imaging of uncommon tumors of the pancreas, Radiol. Clin. North Am., Vol. 40, pp. 1273–1287, 2002.

[37] Stanley, R. J. and Semelka, R. C., Pancreas, In: Computed Body Tomography with MRI Correlation, Lee, J. K. T., Sagel, S. S., Stanley, R. J., and Heiken, J. P., eds., Lippincott Raven, pp. 915–936, 1998.

[38] Sheedy, P. F., II, Stephens, D. H., Hattery, R. R., MacCaty, R. L., and Williamson, B., Jr., Computed Tomography of the Pancreas: Whole Body Computed Tomography, Radiol. Clin. North Am., Vol. 15, No. 3, pp. 349–366, 1977.

[39] Masero, V., Leon-Rojas, J. M., and Moreno, J., Volume reconstruction for health care: A survey of computational methods, Ann. N Y Acad. Sci., Vol. 980, pp. 198–211, 2000.

[40] Udupa, J. K., Three-dimensional visualization and analysis methodologies: A current perspective, Radiographics, Vol. 19, No. 3, pp. 783–806, 1999.

[41] Gonzalez, R. C. and Woods, R. E., (Eds.), Digital Image Processing, 2nd edn., Computer Science Press, Prentice Hall, NJ, 2002.

[42] Kobashi, M. and Shapiro, L. G., Knowledge-based organ identification from CT images, Patt. Recogn., Vol. 28, No. 4, pp. 475–491, 1995.

[43] Dawant, B. M. and Zijdenbos, A. P., Image segmentation, In: Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis, Fitzpatrick, J. M. and Sonka, M., eds., SPIE, pp. 71–127, 2000.

[44] Schiemann, T., Michael, B., Tiede, U., and Hohne, K. H., Interactive 3D-segmentation, SPIE, Vol. 1808, pp. 376–383, 1992.

[45] Ikeda, M., Shigeki, I., Ishigaki, T., and Yamauchi, K., Evaluation of a neural network classifier for pancreatic masses based on CT findings, Comput. Med Imaging Graphics, Vol. 21, No. 3, pp. 175–183, 1997.

[46] Clarke, L. P., Velthuizen, R. P., Camacho, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., and Silbiger, M. L., Review of MRI segmentation: Methods and applications, Magn. Reson. Imaging, Vol. 13, No. 3, pp. 343–368, 1995.

[47] Bensaid, A. M., Improved Fuzzy Clustering for Pattern Recognition with Applications to Image Segmentation., Ph.D. Dissertation, Department of Computer Science, University of South Florida, 1994.

[48] Bezdek, J. C., Pattern Recognition with Fuzzy Objective Function Algorithm, Plenum Press, New York, 1981.

[49] Bensaid, A. M., Bezdek, J. C., Hall, L. O., and Clarke, L. P., A partially supervised fuzzy c-means algorithm for segmentation of MR images, SPIE, Vol. 1710, pp. 522–528, 1992.

[50] Bensaid, A. M., Hall, L. O., Bezdek, J. C., Clarke, L. P., Silbiger, M. L., Arrington, J. A., and Murtagh, R. F., Validity-guided (re)clustering with application to image segmentation, IEEE Trans. Fuzzy Sys., Vol. 4, No. 2, pp. 112–123, 1996.

[51] Clark, M. C., Hall, L. O., Goldgof, D. B., Clarke, L. P., Velthuizen, R. P., and Silbiger, M. S., MRI segmentation using fuzzy clustering techniques, IEEE Eng. Med. Biol. Magazine, Vol. 13, No. 5, pp. 730–742, 1994.

[52] Clarke, L. P., Velthuizen, R. P., Phuphanich, S., Schellenberg, J. D., Arrington, J. A., and Silbiger, M. L., MRI: Stability of three supervised segmentation techniques, Magn. Reson. Imaging, Vol. 11, No. 1, pp. 95–106, 1993.

[53] Vaidyanathan, M., Clarke, L. P., Velthuizen, R. P., Phuphanich, S., Bensaid, A. M., Hall, L. O., Bezdek, J. C., Greenburg, H., Trotti, A., and Silbiger, M., Comparison of supervised MRI segmentation methods for tumor volume determination during therapy, Magn. Reson. Imaging, Vol. 13, No. 5, pp. 719–728, 1995.

[54] Velthuizen, R. P., Clarke, L. P., Phuphanich, S., Hall, L. O., Bensaid, A. M., Arrington, J. A., Greenberg, H. M., and Silbiger, M. L., Unsupervised measurement of brain tumor volume on MR images, J. Magn. Reson. Imaging, Vol. 5, No., 5, pp. 594–605, 1995.

[55] Velthuizen, R. P., Hall, L. O., and Clarke, L. P., An initial investigation of feature extraction with genetic algorithms for fuzzy clustering, Biomed. Eng., Appl., Basis Commun., Vol. 8, No. 6, pp. 496–517, 1996.

[56] Velthuizen, R. P. and Gangadharan, D., Mammographic mass classification: Initial results, In: SPIE Medical Imaging Conference, San Diego, CA, February 12–18, 2000.

[57] Li, L., Zheng, Y., Kallergi, M., and Clark, R. A., Improved method for automatic identification of lung regions on chest radiographs, Acad. Radiol., Vol. 8, pp. 629–638, 2001.

[58] Kallergi, M., Carney, G., and Gaviria, J., Evaluating the performance of detection algorithms in digital mammography, Med. Phy., Vol. 26, No. 2, pp. 267–275, 1999.

[59] Kallergi, M., Clark, R. A., and Clarke, L. P., Medical image databases for CAD applications in digital mammography: Design issues, In: Medical Informatics Europe '97, Pappas, C., Maglaveras, N., and Scherrer, J. R., eds., IOS Press, Amsterdam, pp. 601–605, 1997.

[60] Harrell, F. E., Lee, K. L., and Mark, D. B., Tutorial in biostatistics. Multivariate prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Stat. Med., Vol. 15, pp. 361–387, 1996.

[61] Roe, C. A. and Metz, C. E., Dorfman–Berbaum–Metz method for statistical analysis of multireader, multimodality receiver operating

characteristic data: Validation with computer simulation, Acad. Radiol., Vol. 4, pp. 298–303, 1997.

[62] Li, L., Zheng, Y., Kallergi, M., and Clark, R. A., Improved method for automatic identification of lung regions on chest radiographs, Acad. Radiol., Vol. 8, pp. 629–638, 2001.

[63] Pavlidis, T., Algorithms for Graphics and Image Processing, Computer Science Press, Rockville, MD, 1982.

[64] Greenberg, S., Aladjem, M., Kogan, D., and Dimitrov, I., Fingerprint image enhancement using filtering techniques, In: International Conference on Pattern Recognition, Vol. 3, Barcelona, Spain, Sept. 3–8, 2000.

[65] Heine, J. J., Deans, S. R., Cullers, D. K., Stauduhar, R., and Clarke, L. P., Multiresolution statistical analysis of high resolution digital mammograms, IEEE Trans. Med. Imaging, Vol. 16, pp. 503–15, 1997.

[66] Weaver, J. B., Xu, Y. S., Healy, D. M., Jr., and Cromwell, L. D., Filtering noise from images with wavelet transforms, Magn. Reson. Med., Vol. 21, No. 2, pp. 288–295, 1991.

[67] Hall, L. O., Bensaid, A. M., Clarke, L. P., Velthuizen, R. P., Silbiger, M. L., and Bezdek, J., A Comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain, IEEE Trans. Neural Networks, Vol. 3, No. 5, pp. 672–682, 1992.

[68] Phillips, W. E., Velthuizen, R. P., Phuphanich, S., Hall, L. O., Clarke, L. P., and Silbiger, M. L., Application of fuzzy c-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme, Magn. Reson. Imaging, Vol. 13, No. 2, pp. 277–290, 1995.

[69] Kallergi, M., Gavrielides, M. A., He, L., Berman, C. G., Kim, J. J., and Clark, R. A., A simulation model of mammographic calcifications based on the ACR BIRADS, Acad. Radiol., Vol. 5, pp. 670–679, 1998.

[70] Kallergi, M., He, L., Gavrielides, M., Heine, J. J., and Clarke, L. P., Resolution effects on the morphology of calcifications in digital mammograms, In: Proceedings of VIII Mediterranean Conference on Medical and

Biological Engineering and Computing, Medicon' 98, Lemesos, Cyprus, (June 14–17, 1998), CD-ROM, ISBN 9963-607-13-6.

[71] Zhang, Y. J., A review of recent evaluation methods for image segmentation, In: Proceedings of International Symposium on Signal Processing and its Applications, Malaysia, August 13–16, 2001.

[72] Zhang, Y. J., A survey on evaluation methods for image segmentation, Patt. Recogn., Vol. 29, No. 8, pp. 1335–1346, 1996.

[73] Gerig, G., Jomier, M., and Chakos, M., Valmet: A new validation tool for assessing and improving 3D object segmentation, MICCAI, Vol. 2208, pp. 516–528, 2001.

[74] Chalana, V. and Kim, Y., A methodology for evaluation of boundary detection algorithms on medical images, IEEE Trans. Med. Imaging, Vol. 16, No. 5, pp. 642–652, 1997.

[75] Kelemen, A., Székely, G., and Gerig, G., Elastic model-based segmentation of 3-D neuroradiological data sets, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 828–839, 1999.

[76] Motulsky, H., Intuitive Biostatistics, Oxford University Press, USA, 1995.

[77] Mould, R. F., Introductory Medical Statistics, 3rd edn., Institute of Physics Publishing, Bristol, 1998.

[78] Metz, C. E., ROC methodology in radiologic imaging, Invest. Radiol., Vol. 21, pp. 720–733, 1986.

Chapter 5

Computerized Analysis and Vasodilation Parameterization in Flow-Mediated Dilation Tests from Ultrasonic Image Sequences

Alejandro F. Frangi,[1,2] *Martín Laclaustra,*[3] *and Jian Yang*[1]

5.1 Introduction

Assessment and characterization of endothelial function in the diagnosis of cardiovascular diseases is a current clinical research topic [1, 2]. The endothelium shows measurable responses to flow changes [3, 4], and flow-mediated dilation (FMD) may therefore be used for assessing endothelial health; B-mode ultrasonography (US) is a cheap and noninvasive way to estimate this dilation response [5]. However, complementary computerized image analysis techniques are still very desirable to give accuracy and objectivity to the measurements [1].

Several methods based on the detection of edges of the arterial wall have been proposed over the last 10 years. The first studies used a tedious manual procedure [5], which had a high interobserver variability [6]. Some interactive methods tried to reduce this variability by attracting manually drawn contours to image features, like the maximum image gradient, where the vessel wall is assumed to be located [7–9]. Some more recent efforts are focused on dynamic programming or deformable models [10–19] and on neural networks [20].

[1] Computer Vision Lab, Aragon Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain

[2] Department of Technology, Pompeu Fabra University, Barcelona, Spain

[3] Lozano Blesa University Clinical Hospital, Aragon Institute of Health Sciences, Zaragoza, Spain

All these methods present some common limitations. First, edge detection techniques are undermined by important error sources like speckle noise or the varying image quality typical of US sequences. Second, most methods require expert intervention to manually guide or correct the measurements thus being prone to introduce operator-dependent variability. Also, almost no method performs motion compensation to correct for patient and probe position changes. This could easily lead to measuring arterial dilation using wrong anatomical correspondences. Temporal continuity is another aspect that has not been exploited enough in previous work. Two consecutive frames have a high correlation, and only Newey and Nassiri [20] and Fan *et al.* [15] take advantage of this feature during edge detection. Finally, there is a general lack of large-scale validation studies in most of the techniques presented so far.

In this chapter a method is proposed that is based on a global strategy to quantify flow-mediated vasodilation. We model interframe arterial vasodilation as a superposition of a rigid motion (translation and rotation) and a scaling factor normal to the artery. Rigid motion can be interpreted as a global compensation for patient and probe movements. The scaling factor explains arterial vasodilation. The US sequence is analyzed in two phases using image registration to recover both rigid motion and vasodilation. Image registration uses normalized mutual information [21] and a multiresolution framework [22]. Temporal continuity of registration parameters along the sequence is enforced with a recursive filter. Application of constraints on the vasodilation dynamics is a natural step, since the dilation process is known to be a gradual and continuous physiological phenomenon.

Once a vasodilation curve is obtained, clinical measurementes must be extracted from it. Classically, FMD is quantified by measuring the peak vasodilation diameter relative to the basal diameter level, which are usually manually identified in the curve. Automatically extracting these two parameters is not trivial (e.g., a mere mean of several basal frames and a simple search for a maximum in the curve) given that the curve may also include artifacts. We examined the use of a robust principal component analisys to derive intuitive morphological parameters from the curves and relate them to classical FMD indexes and cardiovascular (CVD) risk factors.

The chapter is organized as follows: Section 5.2 describes the system for image acquisition and the protocol for a typical FMD study. It also describes the population used to evaluate our technique. Section 5.3 introduces the proposed method to assess FMD. The validation of the technique is reported

in section 5.4. In section 5.5 a novel parameterization of the vasodilation curve is introduced, and correlation analyses are presented that relate these new parameters to CVD risk factors and classical FMD parameters. In section 5.6 the results are discussed and some concluding remarks are made in section 5.7.

5.2 Materials

5.2.1 Subjects

A total of 195 sequences of varying image quality were studied, corresponding to 195 male volunteers of the Spanish Army (age range, 34–36 years). This sample is part of the AGEMZA Study, a national cohort study of cardiovascular risk factors in young adults and includes subjects with a wide range of clinical characteristics (body weight: 62.3–111.8 Kg; body mass index: 20.59–35.36 Kg/m^2; hypertension: 9%; hypercholesterolemia: 20%; smokers: 24%).

5.2.2 Image Acquisition

Image acquisition was carried out at the Lozano Blesa University Hospital (Zaragoza, Spain). The echographic probe was positioned onto the arm of the patient lying supine on a bed. A silicon gel was used as impedance adapter for better ultrasound wave transmission. The probe, once the correct orientation angle was found, was fixed with a probe holder to the table where the patient's arm lies (Fig. 5.1). Telediastolic images were captured and hold, coincident with the peak of the R wave of the electrocardiogram. A SONOS 4500 (Agilent Technologies, Andover, MA, USA) ultrasound system was used in frequency fusion mode and employing a 5.5–7.5 MHz trapezoidal multifrequency probe. Images were transferred to a frame grabber via a video Y/C link and images were digitized at a resolution of 768 $\times$ 576 pixels.

During the examination, unavoidable movements take place thus changing the relative position between the transducer and the artery. Therefore, expert intervention is sometimes required to control the image quality by readjusting the orientation of the transducer to keep visible borders as sharp as possible. Both, motion artifacts and successive readjustments may induce changes in image quality as well as changes on extraluminal structures along the sequence. All these factors have to be handled appropriately in the postprocessing stage if the computerized analysis has to be used on a routine basis.

Figure 5.1: Experimental setup.

Each sequence has about 1200 frames and a duration of around 20 min, acquiring each second the last telediastolic frame previously hold. This provides a fixed sampling rate irrespective of heart rate, which means a substantial benefit for clinical interpretation, as different stimulus are applied on a time basis along the clinical test. As the dynamics of endothelial vasodilation is much slower than changes happening between cardiac cycles, with this sampling rate, missing information from one heartbeat or using one heartbeat twice does not affect, in practice, the results.

FMD is the vasodilation response to hyperemia after a transitory distal ischemia induced in the forearm using a pneumatic cuff distal to the probe (Fig. 5.1). The dilation mediated by a chemical vasodilator, the nitroglycerin, or nitroglycerin-mediated dilation (NMD), is also registered. Accordingly, five phases of the medical test can be distinguished in each sequence (see Fig. 5.2).

- *Rest baseline (B1)*. Initial rest state preceding distal ischemia. Presents the best image quality in the whole sequence and lasts for about 1 min.

- *Distal ischemia (DI)*. The cuff is inflated and, therefore, the image quality is usually the worst in the sequence. It takes approximately 5 min. This phase ends when the cuff pressure is released.

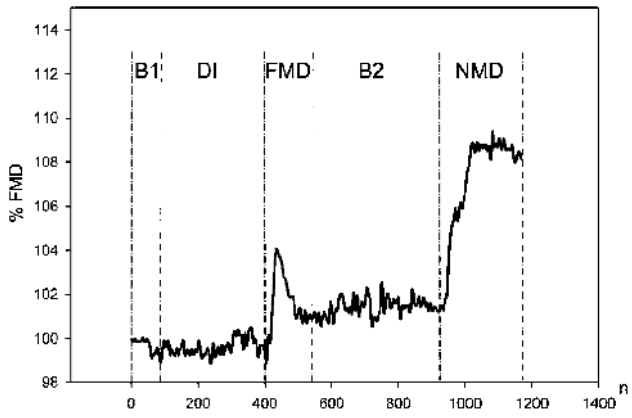


Figure 5.2: A whole typical examination can be divided into several segments along the sequence: rest baseline (B1), distal ischemia (DI), flow-mediated dilation (FMD), post-FMD baseline (B2), and nitroglycerin-mediated dilation (NMD).

- *Flow-mediated dilation (FMD)*. Response to reactive hyperemia. The maximum %FMD of healthy subjects has a mean value around 5%, and it takes place about 60 sec after the cuff is released [5]. At the beginning of this phase, and coincident with the release of the cuff, it is common to have important motion artifacts. The duration of this phase depends on each patient.

- *Post-FMD baseline (B2)*. The artery diameter returns to a steady state. This state does not necessary have to coincide with that of B1.

- *Nitroglycerin-mediated dilation (NMD)*. Response to the sub lingual supply of a fixed dose of this chemical vasodilator, which is made 8 min later than the cuff release.

5.3 Registration-Based FMD Estimation

5.3.1 Algorithm Overview

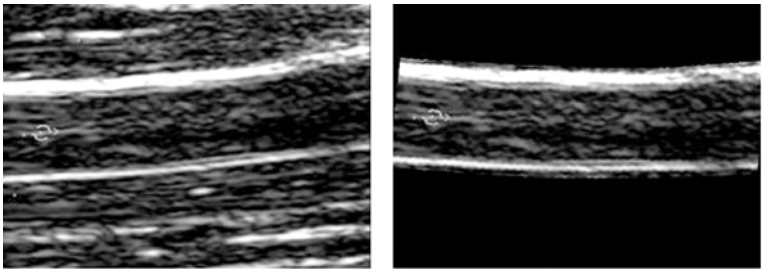
Our technique assumes that the vasodilation that takes place between two frames can be modeled by a constant scaling in the direction normal to the artery. This scale factor is obtained by means of image registration.

Figure 5.3: Overview of the proposed two-stage method: after motion compensation is carried out by recovering a rigid motion model, the vasodilation is measured by computing the scaling factor along the normal to the artery that best matches the two analyzed frames.

A reference frame is selected from the beginning of the sequence. All the other frames are registered to this reference frame. Changes in the relative position between the patient and the transducer are quite common during a whole examination, which may take up to 20 min. To elude wrong anatomical correspondences, motion compensation becomes necessary, and a rigid image registration technique is used to this end.

Structures surrounding the artery in the image may be important to resolve potential ambiguities in the longitudinal alignment between two frames, which occur because of the lineal nature of the arterial walls. On the other hand, extraluminal structures introduce artifacts when measuring the vasodilation since they do not necessarily deform in the same way as the artery does. Therefore, they should be taken into account when retrieving the global rigid motion information of the model, while arterial vasodilation estimation should only consider the artery deformation.

Our technique proceeds in two phases as summarized in Fig. 5.3: motion compensation and dilation assessment. The first phase uses the original frames and rigid image registration to recover a rigid motion model. Translation and rotation parameters are used to initialize the subsequent phase of vasodilation estimation. This second stage employs an affine registration model. To avoid artifacts when measuring arterial vasodilation it is convenient to remove background extra luminal structures by padding them out from the reference frame. Preprocessing of this frame also requires repositioning it so that the artery is normal to the scaling direction, that is to say, aligned with the horizontal axis, since our model searches for a vertical scaling factor (see Fig. 5.4). Both operations are performed manually on the reference frame. Manual masking only requires to roughly draw two lines in the reference frame and repositioning, to align a line with the direction of the artery, a process that is simple and takes only a few seconds per image sequence.



(a) (b)

Figure 5.4: Preprocessing applied to the reference frame before the phase of vasodilation assessment. (a) Original reference frame and (b) the reference frame after alignment to the horizontal axis and padding out of background structures.

Temporal continuity is enforced in both phases by means of recursive filters in the registration parameter space prior to registering each new frame.

In the next two subsections the registration algorithm and its initialization, enforcing temporal continuity, are discussed.

5.3.2 Registration Algorithm

5.3.2.1 Motion and Vasodilation Models

Registering image B onto image A requires finding a transformation $T(\cdot)$ that maps B into A by maximizing a registration measure $M(\cdot)$ as indicated in Eq. (5.1). The similarity measure is computed over all points, P, of the overlap region of both images

$$M(A(P), B(T(P))) \tag{5.1}$$

Our motion model between the original (x, y) and transformed (x', y') coordinates is a rigid transformation of the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{5.2}$$

As we are imaging only a small and roughly straight vessel segment, the vasodilation can be assumed to be normal to the artery and, therefore, it can be modeled by only a scaling factor in that direction. Then, the vasodilation model

Table 5.1: Different similarity measures

SSD	Sum of squared differences
CC	Cross correlation
GCC	Gradient image cross correlation
JE	Joint entropy
MI	Mutual information
NMI	Normalized mutual information

is a similarity transformation with four degrees of freedom:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -s_y \cdot \sin\theta \\ \sin\theta & s_y \cdot \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{5.3}$$

5.3.2.2 Registration Measure

Several registration measures have been traditionally used in medical image matching and they main ones are listed in Table 5.1.

Among the different similarity measures, normalized mutual information (NMI) is selected in this work because of its low sensitivity to the size of the overlap region [21] and higher accuracy (see section 5.4.2.2). This measure, is defined as

$$\mathrm{NMI}(A, B) = \frac{H(A) + H(B)}{H(A, B)} \tag{5.4}$$

where $H(A)$ is the entropy of image A defined as

$$H(A) = -\sum_i p_i \log p_i \tag{5.5}$$

and $H(A, B)$ is the joint entropy between images A and B defined as

$$H(A, B) = -\sum_{i,j} p_{i,j} \log p_{i,j} \tag{5.6}$$

The entropies are computed from the image histograms where p_i is an approximation of the probability of occurrence of intensity value i. Similarly, the joint entropy is computed from the joint histogram where $p_{i,j}$ is the approximation of the probability of the occurrence of corresponding intensity pairs (i, j). Linear interpolation is used to obtain intensities in noninteger pixel values to build the joint histograms.

Table 5.2: Summary of registration parameters

Parameter	Symbol	Value
Registration measure	M	See Table 5.1
No. of bins in joint histogram discretization	b	64
Gaussian kernel width for preblurring	σ_b	1
Resampling ratio	ρ	1.5
No. of resolution levels	r	3
Interpolation scheme	—	Bilinear

5.3.2.3 Optimization Algorithm

A multiresolution framework proposed by Studholme *et al.* [22] is employed to recover the optimal transformation. The image is iteratively subsampled by a factor of two to build a multiresolution image pyramid. The registration problem is solved at each pyramid level in a coarse to fine fashion. The registration parameters found at each level are used as starting estimates for the parameters at the next level.

5.3.2.4 Summary of Registration Parameters

In Table 5.2, a summary of the parameters of the registration algorithm is provided. To compute the several registration metrics of Table 5.1, the joint histogram is discretized using 64×64 bins. Prior to image registration, the images are prefiltered with a Gaussian kernel of $\sigma_b = 1$ pixel and the images are resampled to a new pixel size of 1.5 with respect to the original size. These two steps can help to reduce small-scale noise and to reduce the computational load, and yield seemly registration results. Finally, the optimization strategy proceeds in three resolution levels. Image interpolation is carried out using a bilinear interpolation scheme.

5.3.3 Temporal Continuity

In order to perform motion compensation and vasodilation assessment, it is convenient to introduce prior knowledge about the smooth nature of the arterial vasodilation process. This *a priori* information could be used to filter out sharp transitions in the vasodilation parameter, which arise as a consequence

of registration errors and which are not physiologically plausible. Moreover, these registration errors could easily propagate to the following frames thus invalidating all subsequent measurements.

To avoid error propagation and impose constraints on the vasodilation dynamics, a recursive filter is employed in order to improve the estimation of the initial registration parameters. In the next two subsections, the elaboration of the starting estimates in the motion compensation and in the vasodilation stages are presented.

5.3.3.1 Starting Estimate in Motion Compensation

The motion compensation phase involves three parameters: two translations, t_x and t_y, and a rotation angle, θ. A parameter vector is thus defined by $\mathbf{x} = \{t_x, t_y, \theta\}$. Because of the strong nonstationary behavior of motion artifacts, it is not possible to derive an elaborated linear model of the dynamics the parameter vector. Therefore, a simple first-order auto-regressive model, AR(1), was assumed to predict a suitable initialization for the registration algorithm in the nth frame, $\hat{\mathbf{x}}(n)$. This was done according to

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}(n-1) + \gamma(\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1)) \tag{5.7}$$

where $\mathbf{x}(n)$ denotes the parameter vector output after registering the nth frame. Note that there is some implicit delay between $\hat{\mathbf{x}}(n)$ and $\mathbf{x}(n)$ since they refer to parameter vectors before and after the registration process. Equation (5.7) introduces a systematic inertia to changes in the parameter values through the constant γ. This filtering tries to avoid falling into local minima in the parameter space during registration, which would not be temporally consistent with previous history of arterial motion. On the other hand, it might also slow down the ability to track sudden transitions coming from true motion artifacts. A value of $\gamma = 0.1$ has been empirically shown to be a good compromise between these two competing goals and it was used throughout our experiments.

5.3.3.2 Starting Estimate During Vasodilation Assessment

In this stage we assume that the translation and rotation parameters were correctly recovered at the motion compensation stage. Therefore, only the scale factor, s_y, will be tracked over time using a simplified Kalman filtering scheme [23].

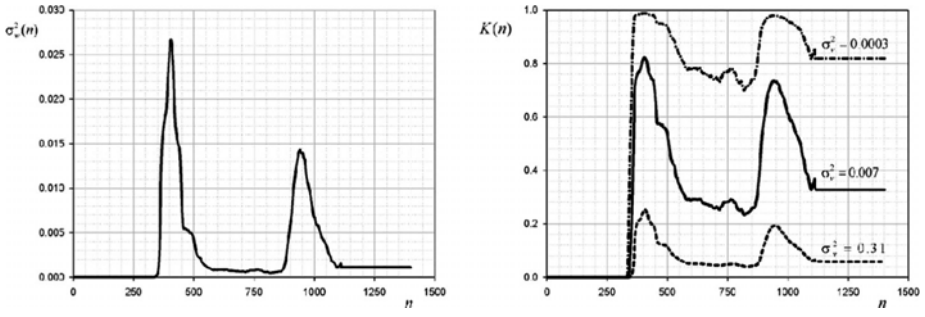


Figure 5.5: Kalman gain. (a) Estimated $\sigma_w(n)$ used for the computation of $K(n)$ during vasodilation assessment. It contains the expected vasodilation dynamics along the sequence. Instants with higher value of $\sigma_w(n)$ correspond to higher uncertainty about the chosen dynamic model and, consequently, where it has to be relaxed to accommodate for possibly sudden transitions. (b) Corresponding Kalman gain for three different measurement noise, $\sigma_v(n)$, values corresponding to the minimum, median, and maximum noise levels, respectively, in our sequence database.

Let us assume that $s_y(n)$ can be also modeled as an AR(1) process

$$s_y(n) = \alpha\, s_y(n-1) + w(n) \tag{5.8}$$

where $w(n)$ is white noise with variance σ_w, and $0 < \alpha < 1$ is the coefficient of the AR(1) model, which was chosen as $\alpha = 0.95$. The scaling factor has a non-stationary behavior and, therefore, $\sigma_w(n)$ actually changes widely over time (cf. Fig. 5.5).

Let the measurement model be

$$\tilde{s}_y(n) = s_y(n) + v(n) \tag{5.9}$$

where $\tilde{s}_y(n)$ are noisy measurements of vasodilation at frame n (obtained via image registration), and $v(n)$ is white noise, uncorrelated with $w(n)$. Under the assumptions of this model, it can be shown [24] that the Kalman filter state estimation equation to predict the vasodilation initialization for the registration of the nth frame, $\hat{s}_y(n)$, is

$$\hat{s}_y(n) = \alpha\, \hat{s}_y(n-1) + K(n)\, [\tilde{s}_y(n-1) - \alpha\, \hat{s}_y(n-1)] \tag{5.10}$$

where $K(n)$ is the Kalman filter gain. Owing to the standardized acquisition protocol, the vasodilation time series has a characteristic temporal evolution,

which can be exploited to give an *a priori* estimation of $\sigma_w(n)$. The value of $\sigma_w(n)$ is high when vasodilation is expected and, it is low when no variations in the artery diameter should be found (e.g. at baselines). On the other hand, the observation noise power $\sigma_v(n)$ will be considered constant, and it is estimated from the first 60 sec when vasodilation is known to be zero. The measurement noise is assumed to be stationary as it mainly depends on the image quality, which can be considered uniform over time for a given sequence.

The temporal evolution of $\sigma_w(n)$ is shown in Fig. 5.5. It has been estimated from the analysis of 50 vasodilation curves (from the dataset described in section 5.2) that were free from artifacts and obtained with the computerized method but using a fixed $K(n) = 0.1$ to assess the vasodilation. The plot indicates the average instantaneous power over the 50 realizations. The 60 initial frames are processed with $K(n) = 0$ to calculate σ_v in each sequence, because no vasodilation is expected during this interval and any variation here can be regarded as measurement noise.

The values of $\sigma_w(n)$ and σ_v determine the Kalman gain, $K(n)$, using the following equation [24]:

$$K(n) = \frac{\alpha^2 K(n-1) + \sigma_w^2(n)/\sigma_v^2}{\alpha^2 K(n-1) + \sigma_w^2(n)/\sigma_v^2 + 1} \tag{5.11}$$

5.4 Computerized FMD Analysis Performance

5.4.1 Examples

Image registration between two frames searches for the transformation that puts them into correspondence. To visually illustrate the algorithm performance, four examples are shown in Fig. 5.6. In four sequences, the reference frame has been aligned with the frame showing maximum FMD.

5.4.2 Evaluation

Three properties of the proposed method are analyzed: accuracy (agreement with the gold standard), reproducibility (repeatability), and robustness (degree of automation of the measurement without apparent failure). To evaluate

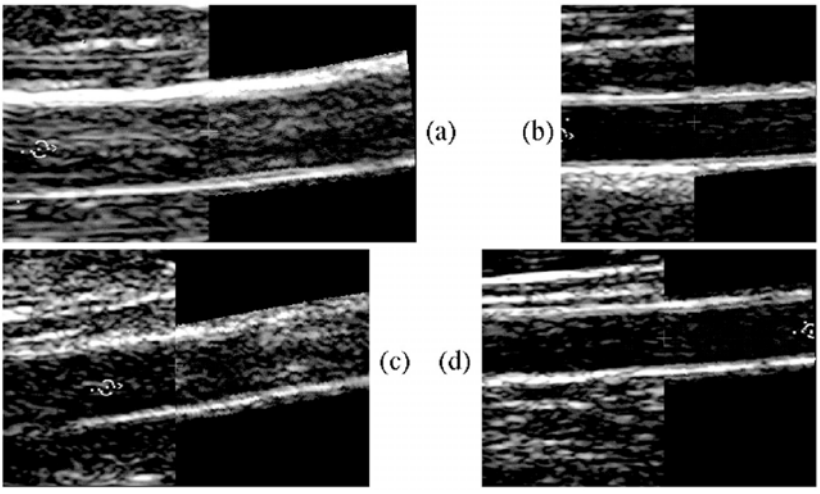


Figure 5.6: Four registration examples of four sequences of different image quality. Each figure shows the flow-mediated dilation frame where maximum vasodilation occurs (left half) registered to the reference B1 frame (right half).

accuracy, a number of expert manual measurements are analyzed and utilized as gold-standard measurements. As well, to get a pattern to compare our method with, accuracy and reproducibility of these manual measurements are also calculated.

We define %FMD as measurement unit for all the sequence values analyzed and represent the relative diameter of a given frame in the sequence to the mean diameter over phase B1 expressed as percentage.

Two statistical methods are used. Firstly, we used Bland–Altman plots [25], a classical method to define limits of agreement between two measurement techniques as indicated by $d \pm 1.96\text{SD}$ where d is the mean difference (bias) and SD is the standard deviation of the differences.

Secondly, we used analysis of variance to estimate the variability (reproducibility) of repeated measurements on every frame. We expressed these also as coefficient of variation (CV), obtained from the mean value ($m_{\%\text{FMD}}$) and the standard deviation ($\text{SD}_{\%\text{FMD}}$) of the %FMD measurements as indicated

$$\text{CV} = \frac{\text{SD}_{\%\text{FMD}}}{m_{\%\text{FMD}}} \tag{5.12}$$

5.4.2.1 Manual Measurements

Manual measurements of arterial diameter were performed in 117 frames corresponding to four sequences of different image quality. Three experts assessed each frame twice in independent sessions. In each sequence several frames were measured: 1 out of 10 in phase B1 (frame number 1, 11 . . . 61) and 1 out of 50 during the rest of the test (frame number 101, 151 . . .). Depending on the duration of the sequence the total number of measured frames was between 28 and 30 per sequence.

Each diameter measurement was obtained by manually fitting a spline to the inner contour of each arterial wall. The diameter was defined as the average distance between both spline curves (see Appendix 5.8). The vasodilation measurements were obtained by dividing the manually obtained diameters by the average diameter over phase B1; the dilation was finally expressed as a percentage, getting %FMD values as defined before.

Gold-standard measurements were derived from these 117 frames. The grand-average of the six diameter measurements done by the three observers is considered the gold-standard arterial diameter estimate for each frame. The gold-standard dilation measurements are obtained by dividing these estimated diameters by the grand-average diameter over phase B1 for each sequence, to get %FMD values. Accuracy, reproducibility, and intra- and interobserver variability of manual measurements were analyzed:

(i) *Accuracy.* Figure 5.7 shows the Bland–Altman plots comparing the intersession average measurement for each observer and the gold-standard measurements. The biases and standard deviation of the differences for

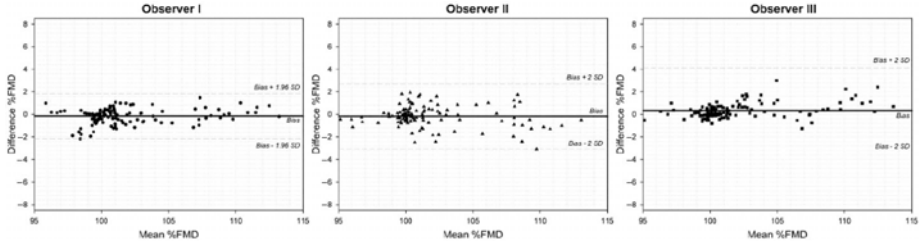


Figure 5.7: Bland–Altman plots comparing the intersession measurement average versus the gold-standard measurements. The horizontal and vertical axes indicate the average %FMD and the difference %FMD, respectively.

Table 5.3: Accuracy of manual measurements

	Obs I	Obs II	Obs III
Bias (%FMD)	−0.16	−0.18	0.34
SD_a(±%FMD)	0.68	0.94	0.68
SD_w(±%FMD)	0.74	1.14	1.41
SD_c(±%FMD)	0.86	1.24	1.20

Bias and standard deviation of the differences (SD_c), corrected for repeated measurements, between manual and gold-standard %FMD measurements. SD_a and SD_w stand for the SD of the differences of the intersession average and the within-observer variability.

the three observers are given in Table 5.3. Standard deviations are corrected to take into account repeated measurements according to the method proposed by Bland and Altman [26].

(ii) *Reproducibility.* The CV of each group of six measurements is calculated for each one of the 117 manually measured frames. This CV is averaged for all the frames of each one of the four sequences, being considered the CV of the manual measurement for each sequence. These four values are averaged finally, obtaining an overall reproducibility value for manual measurements in our study. The results are shown in Table 5.4.

(iii) *Inter- and intraobserver variability.* Figure 5.8 shows Bland–Altman plots comparing both sessions of each observer. In order to estimate the overall inter- and intraobserver variability of manual measurements (with correction for repeated measurements) we carried out the procedure proposed by Bland and Altman in [27]. To this end, a two-way Analysis of Variance (ANOVA) with repeated measurements was performed

Table 5.4: Reproducibility of manual and automated measurements

	CV ($m \pm$ SD)				
	Seq A	Seq B	Seq C	Seq D	Overall
Manual (%)	0.95 ± 0.5	1.20 ± 0.4	0.71 ± 0.6	1.35 ± 0.6	1.04 ± 0.6
Computerized (%)	0.23 ± 0.1	0.26 ± 0.1	0.32 ± 0.3	0.84 ± 0.4	0.40 ± 0.3

Mean and SD of CV (%) measured with respect to %FMD value.

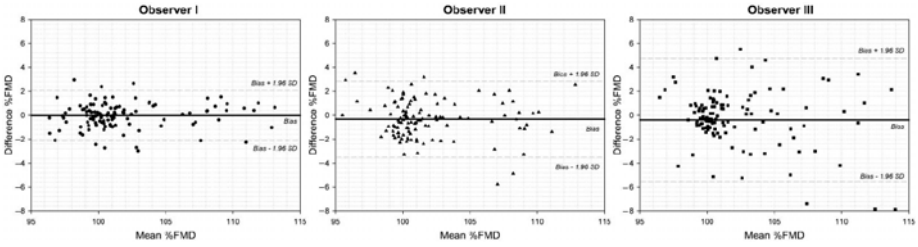


Figure 5.8: Bland–Altman plots comparing the two manual sessions of dilation measurements of each observer. The horizontal and vertical axes indicate the average %FMD and the difference %FMD of the two sessions, respectively.

using Analyse-it v 1.68 (Analyse-it Software Ltd, Leeds, UK). The two-way ANOVA was controlled by observer and measurement frame as fixed factors and by the session number as random factor (Table 5.5). From this analysis, the inter- and intraobserver within-frame %FMD standard deviations were 1.20% and 1.13%, respectively.

5.4.2.2 Computerized Measurements

The scaling factor in the direction normal to the vessel axis that relates each frame to the reference frame constitutes the vasodilation parameter output by the automatic method. As a consequence, the measurements are normalized to the arterial diameter of the reference frame. This normalization is different from that of the gold-standard dilation measurements, which, as described before, were normalized for each sequence to the grand-average diameter over phase B1. To make the computerized measurements comparable to the gold standard,

Table 5.5: A two-way ANOVA of manual measurements of %FMD

Source of variation	SSq	DOF	MSq	F	p
Frame	9329.1	116	80.423	62.82	<0.0001
Observer	19.4	2	9.708	7.58	0.0006
Observer × Frame	359.7	232	1.551	1.21	0.0529
Session	449.4	351	1.280		
Total	10157.6	701			

SSq: Sum of squares; DOF: degrees of freedom; MSq: mean squares; F: F of Snedecor; p: Snedecor test significance.

Table 5.6: Comparison between different similarity metrics

	NMI	MI	GCC	JE	CC	SSD
Bias (%FMD)	+0.05	+0.11	+0.25	−1.00	+1.03	+1.68
SD (± %FMD)	1.05	1.08	2.02	2.49	2.55	3.92

Bias and difference SD in the comparison between the gold standard measures and the automatic dilation obtained with different similarity measures. Values reported correspond to %FMD values. NMI: Normalized mutual information; MI: mutual information; GCC: gradient image cross correlation; JE: joint entropy; CC: cross correlation; SSD: sum of squared differences.

a new normalization of the former measurements is necessary. To this end, the values measured at each frame are divided by the average values over all measurements of phase B1, and are multiplied by a factor 100 to obtain %FMD values.

(i) *Choosing a similarity measure.* Several similarity measures traditionally used in image registration were compared to select the most appropriate one. Thus the four sequences where gold-standard measurements were available were processed using the six similarity measures introduced in Table 5.1. Finally gold-standard vasodilations were compared to the automated vasodilations computed using each registration measure. Table 5.6 indicates that NMI yields the most accurate estimates although the results are only marginally better than using MI. NMI is therefore the similarity measure selected.

(ii) *Accuracy.* Figure 5.9 shows a Bland–Altman plot comparing the automated versus the gold-standard measurements. The SD of the differences is 1.05%. The dilation curves obtained by the proposed method are superimposed to the gold-standard measurements in Fig. 5.10 where we also include the 95% confidence interval of the gold-standard measurements for comparison [26].

(iii) *Robustness.* The whole set of 195 sequences were processed with the proposed method (more than 280,000 frames). The overall result was ranked according to the ability to recover the clinically relevant information from the corresponding vasodilation curve. The results were classified as good, useful, and bad, depending on the amount and severity of the artifacts present in the curve. When, in the opinion of an expert

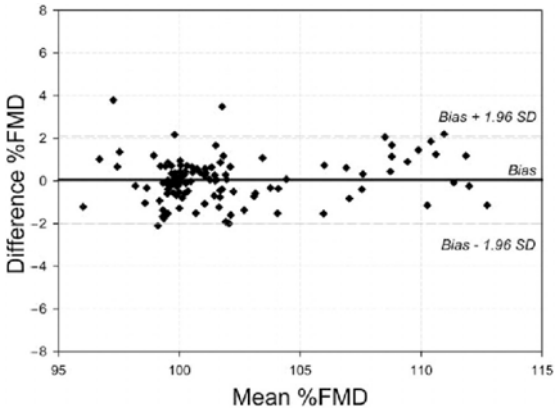


Figure 5.9: Bland–Altman plot comparing the automatic measurements (using normalized mutual information as similarity measure) versus the gold standard. The horizontal and vertical axes indicate the average %FMD and the difference %FMD of the automatic measurements and the gold-standard measurements, respectively.

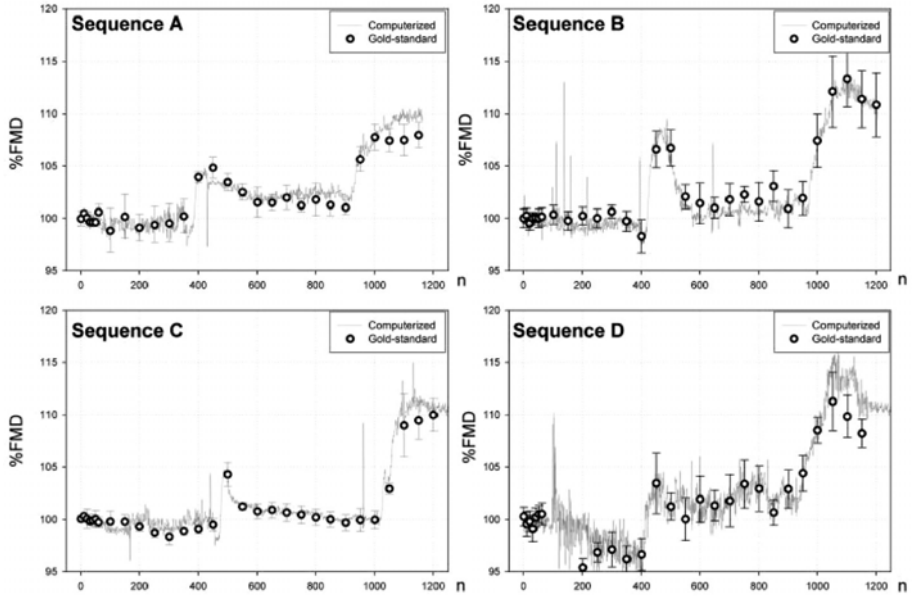


Figure 5.10: FMD curves obtained by the proposed automated method (–) and by the gold-standard measurements (•). Error bars show the 95% confidence interval of the gold-standard measurements for comparison.

sonographer, there were no evident artifacts in the vasodilation curve, the result was scored as good (77.3% of sequences). Artifacts considered were, for instance, lack of convergence or unusual vasodilation evolution. A vasodilation curve was ranked as useful (5.2% of sequences) when artifacts appear only in the DI phase (Fig. 5.2), where no medical information is to be extracted, and therefore, it would still be possible to get clinical information from the other phases. When artifacts appeared in any of the other phases, from where clinical information should be derived, the result was ranked as bad (17.5% of sequences). The vasodilation curves were extracted in a fully automatic fashion with the preprocessing of the reference frame as the only manual intervention from the operator.

(iv) *Reproducibility.* The four sequences with gold-standard measurements were analyzed with the automatic method in six independent runs. Each time a different reference frame was randomly chosen from within phase B1 and it was manually preprocessed (horizontal repositioning of the vessel and removal of extra luminal structures). The CV was computed using as a basis the six dilation measurements for each frame of each sequence. Subsequently, the mean CV in each sequence was obtained by averaging the CV values of the frames where manual measurements were also carried out. These four values are presented in Table 5.4.

5.5 Parameterizing the Vasodilation Response

In the last few years there has been a growing interest in understanding the link between endothelial function and several aspects of cardiovascular diseases (CVD). It is known that impaired endothelial function is associated with a number of disease states, including CVD and its major risk factors [28–31]. Also, endothelial dysfunction seems to precede by many years other more manifest symptoms and may itself be a potentially modifiable CVD risk factor. Therefore, it promises to have not only diagnostic value but also use as an instrument for patient monitoring during treatment.

Once that ongoing research establishes the role and value of FMD in clinical practice, and that computerized tools like the one presented in this chapter

become extensively validated and accepted in clinical protocols, it will be possible to move one step further and use the information contained in FMD curves to derive new clinical indexes of endothelial function. In this section we provide our initial experience in this direction and suggest a possible approach to parameterize FMD curves.

In most clinical papers, FMD is classically measured according to the expression

$$\Delta\mathrm{FMD}_c = \frac{\oslash_{\max} - \oslash_{\mathrm{basal}}}{\oslash_{\mathrm{basal}}} \times 100\% \tag{5.13}$$

where $\oslash_{\max}$ and $\oslash_{\mathrm{basal}}$ are the maximal and basal arterial diameters, respectively. One of the problems generally not discussed in the clinical literature is that manually measuring these diameters can be cumbersome without access to the whole dilation curve. Estimation of $\oslash_{\max}$ by visual inspection of the ultrasound image sequence is time-consuming and subjective, given the small magnitude of the measured dilations. However, using a computerized technique like the one presented here can not only simplify this issue but also provide richer information about the dynamics of the dilation process.

There is still a problem on how to summarize the information present in FMD curves with only a few informative and intuitive parameters. To this end we propose the use of principal component analysis.

5.5.1 Robust EigenFMD and EigenD Modes

Principal component analysis (PCA) [32] is a classical statistical tool to analyze multidimensional datasets by determining the dominant axes according to a maximum variance subspace principle. As the underlying mathematical machinery can be shown to be equivalent to an eigenanalysis of the covariance matrix, PCA axes are usually also named eigenmodes. In cases where outliers can be present in the dataset, there exist robust extensions to PCA, which are less sensitive to wrong or imperfect data samples.

In our application, we performed a PCA-based statistical decomposition of FMD curves and explored the use of the projections on the corresponding eigenmodes (mode coefficients) as potential surrogate indexes of classical $\Delta\mathrm{FMD}_c$. To our purpose we will use a robust version of PCA known as ROBPCA [33] to avoid problems derived from potentially imprecise dilation curves in our dataset.

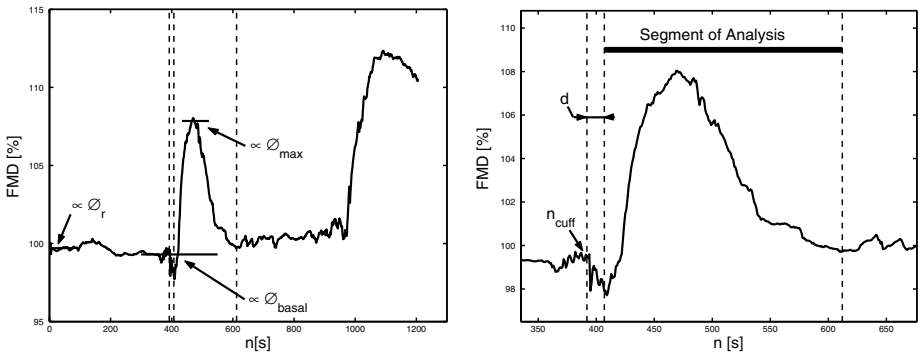


Figure 5.11: Definition of the segment of analysis for principal component analysis (PCA). (a) The original curve with indication of the instant of cuff release (first discontinuous vertical line), and start- ($n_{\text{cuff}} + 15$ sec) and end-points (206 sec later) for PCA (second and third disconstinuous vertical lines); horizontal solid lines indicate the maximum flow-mediated dilation (FMD) and basal FMD levels used for the computation of classical $\Delta\text{FMD}_\text{c}$. Horizontal solid lines indicate the maximum FMD ($\propto \varnothing_{\text{max}}$) and basal FMD levels ($\propto \varnothing_{\text{basal}}$), which are relative to a reference frame ($\propto \varnothing_{\text{r}}$). (b) A zoom into the segment that will undergo PCA. The segment of analysis is indicated with a thick solid line.

To be more specific, let $\mathcal{C}_T(n)$ be an FMD curve (Fig. 5.11(a)). In order to analyze only FMD effects, a segment of this curve, $\mathcal{C}(n)$, was selected for PCA (Fig. 5.11(b)). This segment is defined to have a duration of $D = 206$ sec starting $d = 15$ sec after the release of the cuff pressure. In this way, all segments are synchronized with the FMD onset and have the same duration, which in our acquisition protocol is enough to reach the post-FMD basal level in all subjects without overlapping with the NMD test. We disregard the first 15 sec since that portion of the curve is mostly noisy due to motion artifacts. $\Delta\text{FMD}_\text{c}$ values were also measured by visually selecting the maximal and basal FMD levels. Since the whole curve is relative to the diameter of the reference frame, $\Delta\text{FMD}_\text{c}$ can be calculated directly from this plot. By considering the $\mathcal{C}(n)$ curve as a D-dimensional vector, $\mathbf{c}$, ROBPCA will be applied to the vector set $\{\mathbf{c}_i/\mathcal{C}_i(n_o)\}$, of normalized relative values, where $i = 1 \ldots n$ is the subject index. In our experience, it is convenient to compute the EigenFMD modes with respect to a frame that represents the basal level, hence the division by $\mathcal{C}_i(n_o)$ (cf. next

section). This renormalization corrects for small fluctuations in the baseline due to patient movements in the maneuver of cuff deflation, which themselves can introduce artifactual variation into the statistical analysis. This experiment contains $n = 161$ subjects, whose curves did not have any apparent artifacts in the FMD segment. Figure 5.12 shows the resulting eigenmodes of the relative FMD curves, which will be referred as EigenFMD modes in the sequel.

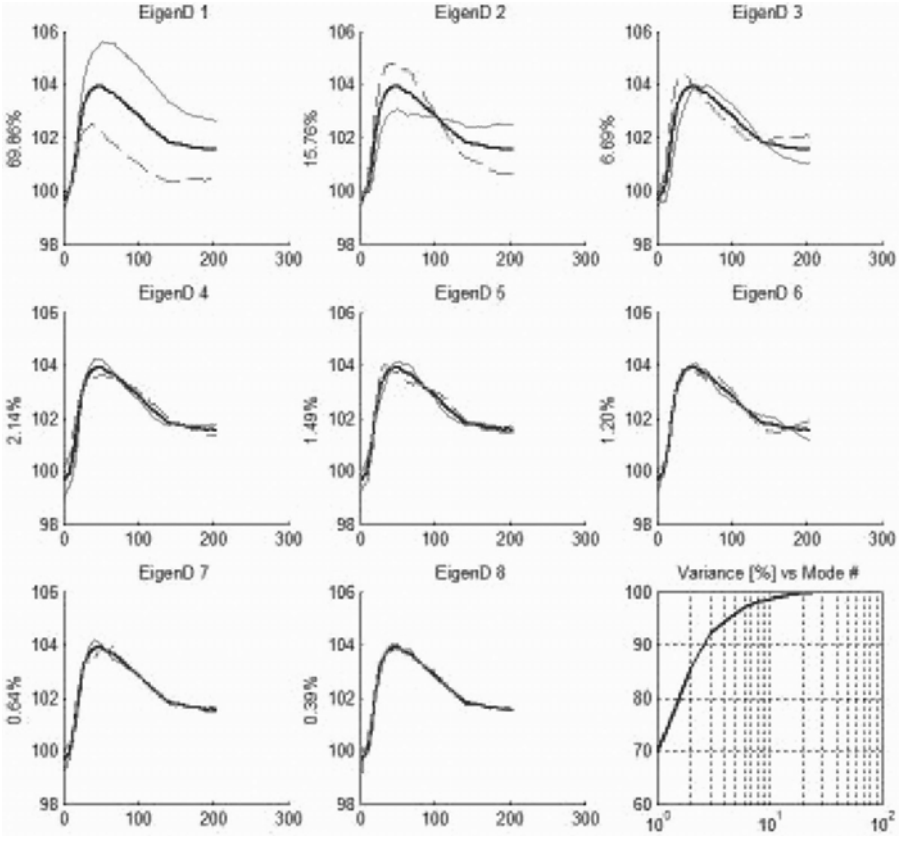


Figure 5.12: Representation of the variation explained by the first eight eigenFMD modes. The average flow-mediated dilation (FMD) curve, and the curves resulting of adding/subtracting the corresponding EigenFMD mode of variation to the mean, weighted by one standard deviation, are plotted in bold line and in thin solid/dashed lines, respectively. The y-axis label of each EigenFMD plot indicates the percentage of variance explained by the corresponding EigenFMD. The lower-right plot shows the accumulated percentage of variance versus the number of EigenFMD modes taken into account. EigenFMD modes were computed using the ROBPCA method.

A similar analysis can be performed by applying the ROBPCA technique to the curves of absolute diameter. This analysis can now reveal correlations between risk factors and the absolute vessel diameter as evidenced by some previous clinical studies. To this end, we manually measured the diameter of the reference frame of each subject, $\oslash_{\mathrm{r}}$, using the technique presented in Appendix 5.8. Subsequently, each vector $\mathbf{c}_i$ was unrelativized according to $\mathbf{c}_i^a = \mathbf{c}_i \cdot \oslash_{\mathrm{r}}$. Figure 5.13 shows the corresponding modes, which are referred as EigenD

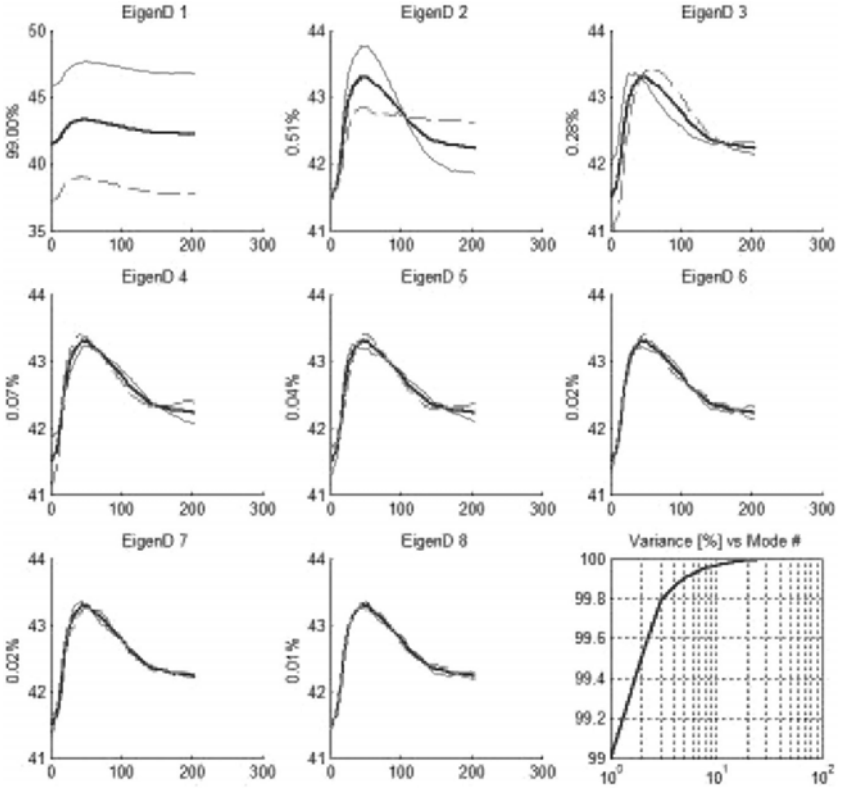


Figure 5.13: Representation of the variation explained by the first eight EigenD modes. The average diameter curve, and the curves resulting of adding/subtracting the corresponding EigenD mode of variation to the mean, weighted by one standard deviation, are plotted in bold line and in thin solid/dashed lines, respectively. The y-axis label of each EigenD plot indicates the percentage of variance explained by the corresponding eigenmode. The lower-right plot shows the accumulated percentage of variance versus the number of EigenD modes taken into account. EigenD modes were computed using the ROBPCA method [33].

modes to emphasize that they correspond to absolute diameter curves. One can observe the fact that the first EigenD mode, which basically represents variations in average vessel diameter, accounts for most of the variability in the dataset and can be regarded as a robust global measure of vessel width. Note that EigenFMD and EigenD modes are similar except for the first mode and the variance attributed to each mode.

5.5.2 Relationship to Classical Indexes

It is important to analyze whether the previous analysis methodology for parameterizing the vessel behavior during the flow-mediated dilation test is linked to other clinical parameters and CVD risk factors traditionally used in the medical literature. Serum lipids, particularly cholesterol and the cholesterol fraction carried by low-density lipoproteins (LDL cholesterol) are recognized as a main causal factor of atherosclerosis [34]. In this disease lipids accumulate in the vessel wall, disturbing the vascular function of delivering sufficient blood flow to the affected territories, which ends with the manifestation of a vascular clinical event like heart attack or stroke. Moreover, knowing patients' lipid levels and modifying them with drugs and diet is the main preventive tool against cardiovascular diseases. From this point of view, cholesterol and LDL cholesterol are considered as risk factors, as higher levels identify individuals with higher risk whereas the cholesterol fraction carried by high-density lipoproteins (HDL cholesterol) is considered as a protective factor. Triglycerides also associate with future cardiovascular disease, and higher levels are present in the metabolic syndrome, which identifies high-risk individuals [34].

Endothelial dysfunction is believed to be a main mechanism through which individuals with unfavorable lipid profiles develop atherosclerosis [35]. As stated before, in the FMD test, the vasodilatory response is identified with endothelial function [1].

5.5.2.1 Sample Population: Clinical and Traditional Variables

A subset of $n = 161$ subjects from the military male population described in section 5.2.1 was analyzed. This population had normal to mildly unfavorable lipid levels. Also, the subset of subjects used in the following analysis corresponds with the one used in computing the EigenFMD and EigenD analyses of the

previous section, and whose FMD curves where obtained with our computerized technique.

Fasting serum samples were obtained from the subjects and were analyzed at the Lozano Blesa University Clinical Hospital (Zaragoza, Spain). Analyses were performed for total cholesterol, triglycerides, and HDL cholesterol by standard enzymatic laboratory techniques. LDL cholesterol was calculated using the Friedewald formula [36,37] in subjects whose triglycerides levels were less than 400 mg/dl.

In each absolute diameter curve, a reference baseline diameter was established before vasodilation, $\oslash_{\text{basal}}$, in a region free from motion artifacts selected just after cuff pressure release (Fig. 5.11(a)). This diameter was similar in most cases to that measured before cuff inflation, $\oslash_{\text{r}}$ (Fig. 5.11(a)). Peak vasodilation was identified in the curve to calculate $\Delta\text{FMD}_{\text{c}}$ (Eq. 5.13).

5.5.2.2 Correlation Analysis

Pearson's linear correlation coefficients were computed between eigenmodes coefficients, from both EigenFMD and EigenD, and serum lipids. Correlations were also computed between ROBPCA-derived indexes and traditional vasodilation curve measurements: flow-mediated dilation ($\Delta\text{FMD}_{\text{c}}$) and baseline diameter ($\oslash_{\text{basal}}$). Pearson's coefficients between traditional curve measurements and serum lipids were used to control the association that could be expected. In all analyses, statistical significance was assumed when $p < 0.05$.

Pearson's linear correlation coefficients between traditional curve measurements and serum lipids are shown in Table 5.7.

Pearson's linear correlation coefficients between EigenFMD and EigenD modes coefficients and serum lipids are shown in Tables 5.8 and 5.9, respectively.

Table 5.7: Pearson's correlation coefficient, r (and p values), between serum lipid levels, and classical $\Delta\text{FMD}_{\text{c}}$ and baseline diameter

Classical index	Cholesterol	Triglycerides	HDL-C	LDL-C
$\Delta\text{FMD}_{\text{c}}$	−0.154 (0.051)	0.035 (0.655)	0.190 (0.016)*	−0.255 (0.001)*
$\oslash_{\text{basal}}$	0.015 (0.853)	0.204 (0.009)*	−0.145 (0.067)	−0.012 (0.884)

Statistically significant correlations are indicated with an asterisk. HDL-C: High-density lipoprotein cholestrol; LDL-C: low-density lipoprotein cholestrol.

Table 5.8: Pearson's correlation coefficient, r (and p values), between serum lipid levels and the first five EigenFMD mode coefficients

Mode #	Cholesterol	Triglycerides	HDL-C	LDL-C
EigenFMD 1	0.053 (0.504)	−0.047 (0.555)	−0.160 (0.043)*	0.132 (0.097)
EigenFMD 2	0.106 (0.180)	0.044 (0.582)	−0.125 (0.115)	0.214 (0.007)*
EigenFMD 3	0.159 (0.044)*	−0.034 (0.672)	−0.027 (0.734)	0.193 (0.015)*
EigenFMD 4	0.016 (0.841)	−0.127 (0.108)	−0.025 (0.750)	0.067 (0.400)
EigenFMD 5	0.044 (0.584)	0.180 (0.022)*	0.020 (0.802)	−0.011 (0.892)

Statistically significant correlations are indicated with an asterisk. HDL-C: High-density lipoprotein cholestrol; LDL-C: low-density lipoprotein cholestrol.

Finally, Pearson's linear correlation coefficients between EigenFMD and EigenD mode coefficients, and classical ΔFMD_c and basal diameter are reported in Tables 5.10 and 5.11.

5.6 Discussion

5.6.1 Computerized FMD Image Analysis

Artery vasodilation assessment is a complex task owing to the poor quality of US image sequences and the small range of the vasodilation that has to be measured. Previous attempts to solve this problem were based on detecting the edges of the arterial wall. These methods have been successful to some extent; however, edge detection in ultrasound is prone to fail due to the presence of speckle

Table 5.9: Pearson's correlation coefficient, r (and p values), between serum lipid levels and the first five EigenD mode coefficients

Mode #	Cholesterol	Triglycerides	HDL-C	LDL-C
EigenD 1	−0.005 (0.949)	−0.217 (0.006)*	0.124 (0.117)	0.032 (0.684)
EigenD 2	−0.109 (0.169)	−0.023 (0.774)	0.096 (0.226)	−0.231 (0.003)*
EigenD 3	−0.072 (0.366)	0.001 (0.992)	−0.128 (0.106)	−0.012 (0.877)
EigenD 4	−0.095 (0.228)	0.058 (0.467)	0.051 (0.517)	−0.145 (0.068)
EigenD 5	0.024 (0.763)	0.147 (0.062)	0.007 (0.934)	−0.027 (0.735)

Statistically significant correlations are indicated with an asterisk. HDL-C: High-density lipoprotein cholestrol; LDL-C: low-density lipoprotein cholestrol.

Table 5.10: Pearson's correlation coefficient, r (and p values), between traditional FMD indexes and the first five EigenFMD mode coefficients

Mode #	ΔFMD_c	$\oslash_{basal}$
EigenFMD 1	−0.786 (<0.001)*	0.118 (0.136)
EigenFMD 2	−0.417 (<0.001)*	0.294 (<0.001)*
EigenFMD 3	−0.319 (<0.001)*	0.076 (0.339)
EigenFMD 4	−0.352 (<0.001)*	−0.190 (0.016)*
EigenFMD 5	0.113 (0.153)	−0.127 (0.108)

Statistically significant correlations are indicated with an asterisk.

noise, poor quality edge definition and acoustic shadows. In our opinion, these techniques are based on low-level features with a poor reliability.

Our method, on the contrary, deals with the images in a more global manner. We model vasodilation as a scaling factor between frames that can be recovered by means of image registration techniques. The effect of low-level artifacts is therefore minimized as the registration measure is computed using all the information present in the whole image, and not just at the edges.

Results obtained with the automated method were better than those measured manually by medical experts. The proposed method presents a negligible bias (0.05 %FMD) whereas the bias of the manual measurements depends on the observer (range −0.16 to +0.34 %FMD). The standard deviation of the differences between the automated and the gold-standard measurements is 1.05 %FMD, which is slightly lower than the intra- and interobserver variabilities of manual measurements (1.13 %FMD and 1.20 %FMD, respectively). From the

Table 5.11: Pearson's correlation coefficient, r (and p values), between traditional FMD indexes and the first five EigenD mode coefficients

Mode #	ΔFMD_c	$\oslash_{basal}$
EigenD 1	0.170 (0.031)*	−0.987 (<0.001)*
EigenD 2	0.322 (<0.001)*	−0.053 (0.507)
EigenD 3	−0.479 (<0.001)*	0.062 (0.437)
EigenD 4	0.479 (<0.001)*	0.005 (0.952)
EigenD 5	0.263 (0.001)*	0.066 (0.409)

Statistically significant correlations are indicated with an asterisk.

dilation CV, the proposed method has also shown to present better reproducibility (CV = 0.40%) than the manual procedure (CV = 1.04%).

The method is reasonably fast. Our experiments were carried out on a PC (Pentium III @ 600 MHz) under RedHat 7.2 Linux operating system. The registration algorithm and the Kalman filtering are both coded in C++ without a thorough code optimization since the implementation of the registration method is a general-purpose software not specifically devised for this application. Under these constraints, the mean execution times per frame are 6.4 sec (SD = 0.8 sec) and 4.0 sec (SD = 1.2 sec) for the first and second phase, respectively. This time also incorporates outputting of progress information. From our experience with the software, we think that these figures could still be cut down substantially by customizing and further optimizing several parts of the code.

The vasodilation model used in this approach has also some potential limitations. Here, dilation is recovered by means of reduced similarity transformation between each frame and the reference one. However, this implicitly assumes that the wall thickness dilates in the same way that the artery does, while it may remain constant or even thin during lumen dilation. The unstable presence of the lumen–intima boundary (LIB) could potentially affect the registration results. Finally, structures stuck to the outer part of the arterial wall may introduce errors in the vasodilation measurements since they make it more difficult to adequately pad the reference frame. The results obtained in this chapter seem to indicate, however, that the vasodilation model outlined in this work is a reasonable simplification.

Motion compensation is necessary to avoid potential sources of bias in the subsequent estimation of vasodilation and to ensure that vasodilation is measured by comparing the same artery segment in two different frames. Nevertheless, stable motion references are required to succeed in motion recovery and avoiding indetermination of the correct alignment in the longitudinal direction of the artery. Moreover, only 2-D information is available in the image to correct a problem that is intrinsically 3-D in nature.

Another advantage of motion compensation is that it makes unnecessary the manual [12, 13] or automatic [15] tracking of a region of interest (ROI) in the image sequence. This ROI tracking is required for the edge detection of some of the methods proposed in the literature. Our technique requires a simple preprocessing of only the reference frame. The interaction required is minimal (only rough delineation of two lines) and introduces a small variability (it is included in the CV of 0.40% obtained in the reproducibility study).

The initialization of the registration algorithm is a very important aspect. This initial transformation should fall inside the capture range of the algorithm [22] whose size depends on many factors, and its determination is not possible *a priori*. Some of these factors are the image quality, the line thickness of the arterial walls and the preprocessing made to the images. It is common that some frames appear with poor image quality along the sequence due to patient's motion. One of these frames may probably lead to erroneous registration values. To reduce error propagation Kalman filtering has shown to be very valuable.

Finally, it is important to recall that the proposed tracking strategy depends on a model of mean arterial vasodilation. This model was estimated from a number of training vasodilation curves, which corresponded to young healthy volunteers. Therefore, this model could bias the analysis of sequences coming from a general population or in specific subject groups like old obese patients. This limitation could be overcome by using a larger training set for building the mean vasodilation model or by having several models for different age groups.

5.6.2 FMD Response Eigen Parameterization

As can be concluded from Table 5.7, a weak but significant correlation was found between $\Delta\text{FMD}_\text{c}$ and LDL cholesterol in a damaging way and between $\Delta\text{FMD}_\text{c}$ and HDL cholesterol in a protective way. This had been previously reported when studying dyslipidemic populations with a similar correlation magnitude, association that could be slightly attenuated in our population due to a narrower range of lipids variation. For instance, Kuvin *et al.* [38] found a correlation coefficient of HDL cholesterol and $\Delta\text{FMD}_\text{c}$ of 0.3. The correlation between classical flow-mediated dilation and LDL cholesterol ($r = -0.40$) had been previously reported by Aggoun *et al.* [39] in hypercholesterolemic patients. Other studies report one of these associations but rarely both or with stronger association coefficients [40] and never in natural large populations with normal lipid levels [41].

Triglycerides correlated significantly with baseline diameter, $\oslash_{\text{basal}}$. The presence of wider vessels has been described for high-risk [38] and atherosclerotic [41] individuals.

The sign associated to the coefficients of EigenFMD and EigenD modes is contingent. These coefficients represent the projections of sample vectors along the eigenvectors of a robust covariance matrix. Since the eigenvectors could equally have been selected with reverse directions the sign of the correlation

ratio has to be interpreted in consonance with the eigenmode plots of Figs. 5.12 and 5.13. On the contrary, the effect of risk factors on the direction of variation of the EigenFMD and EigenD curves is not arbitrary and, therefore, they will be discussed in the sequel.

Among the EigenD modes in Fig. 5.13, the first mode, which broadly represents baseline diameter, significantly correlates with triglycerides, while EigenD mode number 2 does with LDL cholesterol. This later mode graphically appears to represent dilation and diameter decay (the lower the LDL cholesterol level, the higher the peaks and the quicker the vessel recovery).

EigenFMD modes are also related to cholesterol fractions. Interestingly, this analysis highlights that each fraction exerts a different influence on the curve shape. While HDL cholesterol covaries with the first mode, which could be interpreted as a classical measure of FMD peak (the higher the HDL cholesterol level, the higher the peaks), LDL cholesterol is significantly associated with the second mode, which has a form similar to EigenD 2 (the higher the LDL cholesterol levels, the lower the peaks and the slower the decays), and EigenD 3, interpretable as response velocity or time-to-peak (the higher LDL-cholesterol values, the later the peaks).

The classical measurement of FMD, $\Delta \mathrm{FMD_c}$, correlates with statistical significance with almost all absolute and relative shown eigenmodes. The mode with the highest correlation is EigenFMD 1 ($r = -0.786$), which one can visually argue that captures the maximum dilation peak. Other modes with correlation coefficients over 0.300 are also present but their visual interpretation is more subtle.

EigenD 1 is almost equivalent to the baseline diameter. This is promising, as this manual measurement could be potentially estimated on an observer-independent basis. Baseline diameter also showed significant correlations with EigenFMD 2 and 4 although the interpretation of this fact is not so evident.

5.7 Conclusions

This chapter presents a new method to assess brachial artery vasodilation in US sequences. This method, based on image registration, minimizes the effect of low-level artifacts. It also incorporates a motion compensation phase, which relieves the operator of manually tracking a region of interest.

The method is accurate (bias $= +0.05\%$, and limits of agreement $\pm$ 2.05 %FMD), has better reproducibility (CV $= 0.40\%$) than manual

measurements (CV = 1.04%), and is robust, yielding clinically relevant information in at least 80% of the sequences in an uncontrolled setting. Finally, the method requires minimal user intervention having limited effect on the reproducibility of the measurements.

In this work we have also introduced a novel parameterization of the vasodilation curve in the FMD test. This parameterization is of applicability in conjunction with the FMD measurement technique provided in this chapter or with any other alternative (e.g. [12–20]). We have shown that this parameterization yields a number of FMD-related indexes that are consistent with the traditional ΔFMD_c coefficient and basically replicates previous findings in the literature linking FMD to serum lipid levels. This parameterization has several advantages over the classical alternative as it captures more comprehensive information of vasodilation and does not require manual selection of the FMD peak or baseline diameters to extract EigenD modes from absolute diameter curves. However, EigenFMD modes still require manual determination of the basal FMD level or diameter to process relative or absolute curves, respectively. More important, EigenFMD and EigenD coefficients bear a graphical interpretation in terms of FMD-peak, time-to-FMD-peak, recovery speed, etc. This can be useful to gain understanding in the relationship between the dynamics of the endothelium dependent vasodilation and the underlying processes related to endothelial dysfunction. Although this method enables such investigations that are impossible with classical measurements of FMD, further clinical research is required to fully validate its utility.

5.8 Appendix: Distance Between Two Splines

Each manual measurement requires fitting of a cubic spline to the edge of each arterial wall. The distance between both splines is the arterial diameter. The line is placed at the inner edge of the media as showed in Fig. 5.14.

The distance is calculated as follows (see Fig. 5.15):

1. The orientation of each spline is calculated by means of a linear regression.
2. The mean orientation of both splines is calculated using the bisector of the two calculated regression lines.

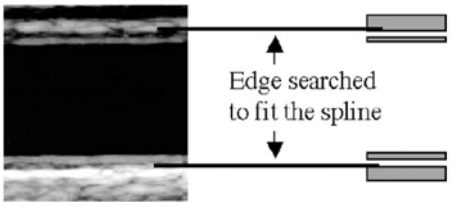


Figure 5.14: Manual fitting of a spline to the arterial wall.

3. Perpendicular lines to this bisector are traced, every 10 pixels, finding the intersection points with the two splines.

4. The average distance between all pairs of points found is the arterial diameter.

5.9 Acknowledgments

We express our gratitude to Dr. D. Rueckert for providing us with his implementation of Studholme's algorithm. We thank M.L. Gimeno, MD, and A.G. Frangi, MD, for providing manual measurements for the evaluation study, and S. Olmos and M. Bossa for their help and discussions. We also thank P. Lamata for his contribution in the initial phases of this work. AFF is supported by a Ramón y Cajal Research Fellowship from the Spanish ministry of Science and Technology.

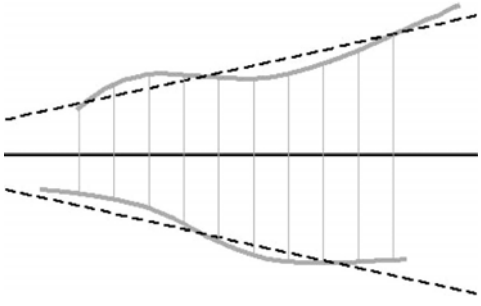


Figure 5.15: Average diameter estimation. The diameter is the mean length of the segments (thin continuous gray lines) perpendicular to the bisector (continuous black line) between the regression lines (dashed black lines) of both splines. Nonparallelism and tortuosity of the splines representing the arterial walls have been exaggerated to clarify the example.

This research was partially supported by grant TIC2002-04495-C02 from the same ministry. The AGEMZA Project is supported by a grant (FIS 99/0600) from the Spanish Ministry of Health and Consumption. The clinical research was also partially supported by a grant of the Diputación General de Aragón (P58/98).

Questions

1. *In section 5.1 it is stated that edge detection approaches to FMD computerized analysis are more dependent on several error sources of ultrasound that registration based approaches. Can you explain why?*

2. *The motion model of section 5.3.2.1 assumes that there is in-plane motion only. Can you comment on this?*

3. *The Kalman filter is causal, which means that its output value is a function of only the inputs that came earlier in time (could also be only later). Comment on the use of non causal tracking strategies like, for instance, noncausal Wiener filtering.*

4. *Derive Eq. (5.11).*

5. *Why do we need to measure the nitroglycerine-mediated dilation (NMD) phase if clinical indexes are only related to the FMD phase?*

Bibliography

[1] Corretti, M. C., Anderson, T. J., Benjamin, E. J., Celermajer, D., Charbonneau, F., Creager, M. A., Deanfield, J., Drexler, H., Gerhard-Herman, M., Herrington, D., Vallance, P., Vita, J., Vogel, R., and International Brachial Artery Reactivity Task Force, Guidelines for the ultrasound assessment of endothelial-dependent flow-mediated vasodilation of the brachial artery: A report of the International Brachial Artery Reactivity Task Force, J. Am. Coll. Cardiol., Vol. 39, pp. 257–265, 2002.

[2] Teragawa, H., Kato, M., Kurokawa, J., Yamagata, T., Matsuura, H., and Chayama, K., Endothelial dysfunction is an independent factor responsible for vasospastic angina, Clin. Sci. (London), Vol. 101, pp. 707–713, 2001.

[3] Cooke, J. P., Rossitch, E., Jr, Andon, N. A., Loscalzo, J., and Dzau, V. J., Flow activates an endothelial potassium channel to release an endogenous nitrovasodilator, J. Clin. Invest., Vol. 88, pp. 1663–1671, 1991.

[4] Sinoway, L. I., Hendrickson, C., Davidson, W. R., Jr, Prophet, S., and Zelis, R., Characteristics of flow-mediated brachial artery vasodilation in human subjects, Circ. Res., Vol. 64, pp. 32–42, 1989.

[5] Celermajer, D. S., Sorensen, K. E., Gooch, V. M., Spiegelhalter, D. J., Miller, O. I., Sullivan, I. D., Lloyd, J. K., and Deanfield, J. E., Noninvasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis, Lancet, Vol. 340, pp. 1111–1115, 1992.

[6] Hardie, K. L., Kinlay, S., Hardy, D. B., Wlodarczyk, J., Silberberg, J. S., and Fletcher, P. J., Reproducibility of brachial ultrasonography and flow-mediated dilatation (FMD) for assessing endothelial function, Aust. N.Z. J. Med., Vol. 27, pp. 649–652, 1997.

[7] Touboul, P. J., Prati, P., Scarabin, P. Y., Adrai, V., Thibout, E., and Ducimetiere, P., Use of monitoring software to improve the measurement of carotid wall thickness by B-mode imaging, J. Hypertens. Suppl., Vol. 10, pp. S37–S41, 1992.

[8] Gariepy, J., Massonneau, M., Levenson, J., Heudes, D., Simon, A., and Groupe de Prevention Cardio-vasculaire en Medecine du Travail, Evidence for in vivo carotid and femoral wall thickening in human hypertension, Hypertension, Vol. 22, pp. 111–118, 1993.

[9] Selzer, R. H., Hodis, H. N., Kwong-Fu, H., Mack, W. J., Lee, P. L., Liu, C. R., and Liu, C. H., Evaluation of computerized edge tracking for quantifying intima-media thickness of the common carotid artery from B-mode ultrasound images, Atherosclerosis, Vol. 111, pp. 1–11, 1994.

[10] Kozick, R., Detecting interfaces on ultrasound images of the carotid artery by dynamic programming, In: IS&T/SPIE Electronic Imaging Symposium, San Jose, CA, Feb 1996, Vol. 2666, pp. 233–241.

[11] Gustavsson, T., Liang, Q., Wendelhag, I., and Wikstrand, J., A dynamic programming procedure for automated ultrasonic measurement of the carotid artery, In: IEEE Computers Cardiology, IEEE Computer Society, pp. 297–300, 1999.

[12] Sonka, M., Liang, W., and Lauer, R. M., Flow-mediated dilatation in brachial arteries: Computer analysis of ultrasound image sequences, CVD Preven., Vol. 1, pp. 147–55, 1998.

[13] Liang, Q., Wendelhag, I., Wikstrand, J., and Gustavsson, T., A multiscale dynamic programming procedure for boundary detection in ultrasonic artery images, IEEE Trans. Med. Imaging, Vol. 19, pp. 127–142, 2000.

[14] Preik, M., Lauer, T., Heiss, C., Tabery, S., Strauer, B. E., and Kelm, M., Automated ultrasonic measurement of human arteries for the determination of endothelial function, Ultraschall. Med., Vol. 21, pp. 195–198, 2000.

[15] Fan, L., Santago, P., Jiang, H., and Herrington, D. M., Ultrasound measurement of brachial flow-mediated vasodilator response, IEEE Trans. Med. Imaging, Vol. 19, pp. 621–631, 2000.

[16] Fan, L., Santago, P., Riley, W., and Herrington, D. M., An adaptive template-matching method and its application to the boundary detection of brachial artery ultrasound scans, Ultrasound Med. Biol., Vol. 27, pp. 399–408, 2001.

[17] Mignotte, M. and Meunier, J., A multiscale optimization approach for the dynamic contour-based boundary detection issue, Comput. Med. Imaging Graph., Vol. 25, pp. 265–275, 2001.

[18] Woodman, R. J., Playford, D. A., Watts, G. F., Cheetham, C., Reed, C., Taylor, R. R., Puddey, I. B., Beilin, L. J., Burke, V., Mori, T. A., and Green, D., Improved analysis of brachial artery ultrasound using a novel edge-detection software system, J. Appl. Physiol., Vol. 91, pp. 929–937, 2001.

[19] Sonka, M., Liang, W., and Lauer, R. M., Automated analysis of brachial ultrasound image sequences: Early detection of cardiovascular disease via surrogates of endothelial function, IEEE Trans. Med. Imaging, Vol. 21, pp. 1271–1279, 2002.

[20] Newey, V. R. and Nassiri, D. K., Online artery diameter measurement in ultrasound images using artificial neural networks, Ultrasound Med. Biol., Vol. 28, pp. 209–216, 2002.

[21] Studholme, C., Hill, D. L. G., and Hawkes, D. J., An overlap invariant entropy measure of 3D medical image alignment, Patt. Recogn., Vol. 32, No. 1, pp. 71–86, 1999.

[22] Studholme, C., Hill, D., and Hawkes, D., Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Med. Phys., Vol. 24, No. 1, pp. 25–35, 1997.

[23] Kalman, R. E., A new approach to linear filtering and prediction problem, Trans. ASME, J. Basic Eng., Vol. 82 (Series D), pp. 35–45, 1960.

[24] Hayes, M., Statistical Digital Signal Processing and Modelling, Wiley, New York, 1996.

[25] Altman, D., Practical Statistical Research, Chapman & Hall, Boca Raton, FL, 1991.

[26] Bland, J. and Altman, D., Measuring agreement in method comparison studies, Stat. Methods Med. Res., Vol. 8, No. 2, pp. 135–160, 1999.

[27] Bland, J. and Altman, D., Statistical methods for assessing agreement between two methods of clinical measurement, Lancet, Vol. 1, No. 8476, pp. 307–310, 1986.

[28] Vita, J. A. and Keaney, J. F., Jr, Endothelial function: A barometer for cardiovascular risk?, Circulation, Vol. 106, No. 6, pp. 640–642, 2002.

[29] Faulx, M. D., Wright, A. T., and Hoit, B. D., Detection of endothelial dysfunction with brachial artery ultrasound scanning, Am. Heart J., Vol. 145, No. 6, pp. 943–951, 2003.

[30] Widlansky, M. E., Gokce, N., Keaney, J. F., Jr, and Vita, J. A., The clinical implications of endothelial dysfunction, J. Am. Coll. Cardiol., Vol. 42, No. 7, pp. 1149–1160, 2003.

[31] Gokce, N., Keaney, J. F., Jr, Hunter, L. M., Watkins, M. T., Nedeljkovic, Z. S., Menzoian, J. O., and Vita, J. A., Predictive value of noninvasively determined endothelial dysfunction for long-term cardiovascular events in patients with peripheral vascular disease, J. Am. Coll. Cardiol., Vol. 41, No. 10, pp. 1769–1775, 2003.

[32] Jolliffe, I., Principal Component Analysis, 2nd edn., Springer Series in Statistics, Springer-Verlag, New York, 2002.

[33] Hubert, M., Rousseeuw, P., and van den Branden, K., ROBPCA: A new approach to robust principal component analysis, Technical Report, Department of Mathematics, Katholieke Universiteit Leuven, 2003.

[34] The National Cholesterol Education Program (NCEP), Executive summary of the third report of The National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III), JAMA, Vol. 285, pp. 2486–2497, 2001.

[35] Ross, R., The pathogenesis of atherosclerosis: A perspective for the 1990s, Nature, Vol. 362, pp. 801–809, 1993.

[36] Friedewald, W. T., Levy, R. I., and Fredrickson, D. S., Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without

use of the preparative ultracentrifuge, Clin. Chem., Vol. 18, pp. 499–502, 1972.

[37] Roberts, W. C., The Friedewald-Levy-Fredrickson formula for calculating low-density lipoprotein cholesterol, the basis for lipid-lowering therapy, Am. J. Cardiol., Vol. 62, pp. 345–346, 1988.

[38] Kuvin, J. T., Patel, A. R., Sidhu, M., Rand, W. M., Sliney, K. A., Pandian, N. G., and Karas, R. H., Relation between high-density lipoprotein cholesterol and peripheral vasomotor function, Am. J. Cardiol., Vol. 92, pp. 275–279, 2003.

[39] Aggoun, Y., Bonnet, D., Sidi, D., Girardet, J. P., Brucker, E., Polak, M., Safar, M. E., and Levy, B. I., Arterial mechanical changes in children with familial hypercholesterolemia, Arterioscler Thromb. Vasc. Biol., Vol. 20, pp. 2070–2075, 2000.

[40] Toikka, J. O., Ahotupa, M., Viikari, J. S., Niinikoski, H., Taskinen, M., Irjala, K., Hartiala, J. J., and Raitakari, O. T., Constantly low HDL-cholesterol concentration relates to endothelial dysfunction and increased in vivo LDL-oxidation in healthy young men, Atherosclerosis, Vol. 147, pp. 133–138, 1999.

[41] Holubkov, R., Karas, R. H., Pepine, C. J., Rickens, C. R., Reichek, N., Rogers, W. J., Sharaf, B. L., Sopko, G., Merz, C. N., Kelsey, S. F., McGorray, S. P., and Reis, S. E., Large brachial artery diameter is associated with angiographic coronary artery disease in women, Am. Heart J., Vol. 143, pp. 802–807, 2002.

Chapter 6

Statistical and Adaptive Approaches for Optimal Segmentation in Medical Images

Shuyu Yang[1] *and Sunanda Mitra*[1]

6.1 Introduction

Image analysis techniques have been broadly used in computer-aided medical analysis and diagnosis in recent years. Computer-aided image analysis is an increasingly popular tool in medical research and practice, especially with the increase of medical images in modality, amount, size, and dimension. Image segmentation, a process that aims at identifying and separating regions of interests from an image, is crucial in many medical applications such as localizing pathological regions, providing objective quantative assessment and monitoring of the onset and progression of the diseases, as well as analysis of anatomical structures.

Generally speaking, segmentation techniques are application specific and nonuniversal. There exists no approach that works best for all types of images. In fact, some approaches work better on one type of image than others, depending on the modality of the image. For example, images acquired by magnetic resonance imaging (MRI) and radiographic X-ray imaging are quite different from retinal images. In the former, the images are represented by intensity variations proportional to radiation absorption or RF signal amplitude mapped into gray-level values, while for the latter, the images are chromatic and generated optically. Numerous segmentation techniques have been developed for gray scale

[1]Department of Electrical and Computer Engineering, Texas Tech University Lubbock, TX 79409-3102

images [1–5], while color image segmentation techniques have been created much later than its gray-level counterpart because of the computational complexity involved with the latter. However, the availability of fast digital processors in recent times allows easy implementations of such complex algorithms. Most of the segmentation techniques applied to gray-level images can also be extended to color images [6, 7].

Clustering is a pattern recognition technique that has been frequently used in image segmentation [8, 9]. Similar to the variety of approaches in image segmentation, there are numerous clustering techniques based on statistics, fuzzy logic [10, 11], neural network, or an integration of these [12]. This chapter applies two recently developed advanced clustering algorithms, namely, deterministic annealing (DA) [13] and adaptive fuzzy leader clustering (AFLC) [14] and compares their performances with other standard well-known algorithms in efficient segmentation of medical images. DA is designed on a statistical frame work, while AFLC has a neural network structure embedded with fuzzy optimization. The performances of these two algorithms have been compared with classical clustering techniques such as k-means [15], and fuzzy C-means (FCM) [16]. These clustering algorithms have been applied to segment a few diverse types of medical images. All operations are performed on images in the spatial domain, i.e., pixel intensity will be used as the only feature. For gray-scale images, such as MRI, the feature will be 1D, while for color images, such as the retinal image, the classification is 3D (red, green, and blue components for each pixel).

The major advantage of using clustering for medical image segmentation is that these unsupervised techniques for data partitioning do not require a training set, which is not easy to find in most clinical datasets. The two clustering techniques, namely AFLC and DA, used in our study to investigate the effectiveness and accuracy of these techniques in medical image segmentation can be considered as optimization processes. Both AFLC and DA do not require an initial guess of the actual number of clusters present in a dataset and thus do not suffer from the instability inherent to traditional and well-known clustering algorithms such as k-means.

Several types of medical images are selected and used as examples of clustering application. The first modality we used is MRI. We compared the segmentation of anatomical structures such as gray matter, white matter, and cerebrospinal fluid from simulated MRI. Pathological segmentation of multiple sclerosis with both simulated and real MRI is also performed. The second imaging

modality involves stereo retinal imaging for evaluating structural damage in the retina. We demonstrated delineation of blood vessels in these images with DA and applied the result to 3-D disparity mapping and segmentation of optic disk/cup. Traditional 2-D segmentation of the optic disk/cup is also presented and compared with manually segmented optic disk/cup by ophthalmologists. The third imaging modality is direct optical (color) imaging used to investigate precancerous lesions in the cervix. Traditionally, an abnormal Pap smear finding is followed by Cervicographic and Colposcopic examination of the cervix by a gynecologist/oncologist to determine the severity of the lesions and identify the location for a biopsy probe. Segmentation of lesions using selected clustering algorithms is carried out for comparison.

The chapter is organized as follows: A brief description of image segmentation is given in section 6.2. Selected clustering techniques, traditional and advanced, are introduced in section 6.3. In section 6.4, results of applying these techniques for precise segmentation of various modalities of medical images are presented. Section 6.5 contains discussions and conclusions.

6.2 Image Segmentation

Image segmentation can be defined as separating the image into similar constituent parts. Given an image I, segmentation of I is a partition P of I into a set of N regions R_n, $n = 1, \ldots, N$, such that $\bigcup_{n=1}^{N} R_n = I$. The separated regions should be homogeneous and meaningful to the application intended. According to Pham *et al.* [1] image segmentation techniques can be classified into several categories, such as thresholding, region growing, classifiers, clustering, Markov random field, artificial neural network, fuzzy logic, deformable models, and atlas-guided approaches. The performance efficiency of each approach, however, varies and is dependent on specific application and image modality. When a practical application is concerned, sometimes integration of these techniques is needed to achieve better performance. A number of review papers on image segmentation in general and specifically on medical image segmentation are already available [1–7, 9]. In this chapter, we have focused on the impact of recent advanced clustering algorithms in precise segmentation of medical images.

Most of the common medical images such as MRI, positron emission tomography (PET), computed tomography (CT), and ultrasound images are

monochromatic. Some of these types of images can be pseudocolored. As mentioned earlier in section 6.1, most of the segmentation techniques developed for gray scale images can be extended to color images. There are quite a few color models [17] that are commonly used in image processing, mainly to comply with color video standards and human perception. RGB (red, green, blue), HSI (hue, saturation, intensity), and CIE L*a*b* are color models that have been frequently used in segmentation. RGB is hardware oriented while HSI and L*a*b* representations are compatible with human visual perception. What is more, perceptual uniformity of the L*a*b* color space is advantageous over RGB and HSI in that the human perception of color difference can be represented as the euclidean distance between color points, a useful property can be used in error functions of some segmentation algorithms. Most color images, such as the ones used in our examples, the color retinal stereo images and the color cervix images, are captured in RGB. If processing in other color space is preferred, color transformation is needed [6, 7].

To verify the efficiency of a segmentation method, segmentation result is compared with the "truth model." Truth models for practical images are often obtained by manual segmentation. In medical imaging, manual segmentation is usually performed by trained medical professionals. However, such truth models are not as accurate since they are prone to subjective variability, and often poor repeatability, leading to some ambiguities in diagnosis. Therefore, as far as validation is concerned, computer-simulated phantoms are preferred. The phantoms are simplified mimics of the real images. Quantative statistics such as the number of misclassified pixels or shape differences can be obtained by comparison segmented result with the phantom in a straightforward manner. Unfortunately, phantoms are not available for all image modalities. In our examples, we used both computer-simulated phantoms (MRI) as well as manual segmentations (stereo retinal image and color cervix images) to validate the clustering algorithms.

6.3 Clustering Methods

Clustering is a natural way for image segmentation since partitions of similar intensity or texture can be seen as different clusters, the same way human beings perceive objects. Let $\mathbf{x}_i$, $i = 1, \ldots, N$, be a sample of the input space, and let $\mathbf{c}_j \subset \mathbf{C}$, $j = 1, \ldots, M$, be one class of a total of M classes. A clustering

algorithm determines the classes **C** and assigns every sample $\mathbf{x}_i$ into one of the classes. For hard clustering, a sample belongs to only one class, meaning $C_k \bigcap C_j = \phi,\ \forall k \neq j$. For fuzzy clustering, a sample can be classified into more than one class with different membership values (a degree of similarity) [11]. The sum of all membership values of one sample is unity. Categorization and summary of most clustering techniques can be found in [12, 18, 19].

k-means (or c-means) and its fuzzy version FCM are two well-known classical clustering algorithms used for image segmentation. A comparative study of k-means and FCM is presented in [20]. Application of these algorithms and their variations on image segmentation can be found in [21–25].

The clustering techniques discussed below, namely, AFLC, DA, k-means, and FCM, can be regarded as optimization processes that seek to reduce misclassification by minimizing specific cost functions or system energy functions. Contrary to the classical Bayesian classifier that needs training, these clustering techniques are unsupervised. The complexity of these algorithms, however, varies. k-means and FCM are relatively simple and easier to implement but not as effective when compared to DA and AFLC, as will be demonstrated in the next section. The main problem inherent to both k-means and FCM is that the initial guess of the actual number of clusters present in a dataset is crucial to the convergence of the algorithms.

6.3.1 Adaptive Fuzzy Leader Clustering

AFLC is an integrated neural-fuzzy clustering algorithm that can be used to learn cluster structure embedded in complex datasets in a self-organizing manner. The algorithm has a two-layer structure as illustrated in Fig. 6.1. The first layer uses a self-organizing neural network similar to ART1 [26–28] to find hard clusters. Let C be the current number of centroids and $\mathbf{v}_i (i = 1, \ldots, C)$ representing the centroids. When a new sample $\mathbf{x}_k$ comes in, it is normalized and then initially classified into the cluster on which it has the largest projection (a winner-take-all or MAXNET [29, 30] learning rule):

$$y_{i^*} = \max_i \{y_i\} = \max_i \left\{ \sum_{j=1}^{n} b_{ij} \overline{x}_{kj} \right\} \quad i = 1, \ldots, C \tag{6.1}$$

$$b_{ij} = \frac{v_{ij}}{\|\mathbf{v}_i\|} \quad \overline{x}_{kj} = \frac{x_{kj}}{\|\mathbf{x}_k\|}$$

where i^* is the index of the winning centroid.

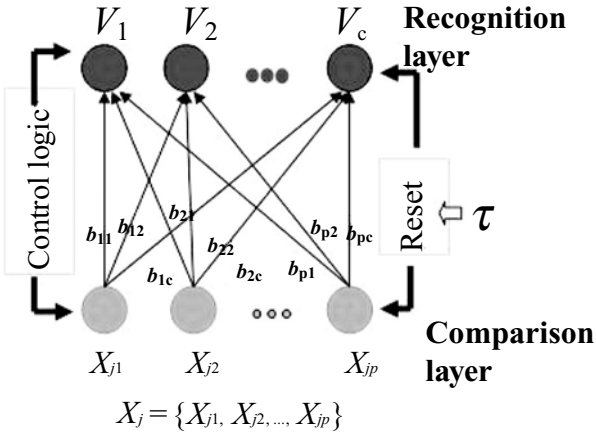


Figure 6.1: Adaptive fuzzy leader clustering (AFLC) structure.

The second layer serves as a verification process. By verifying the initial sample recognition through a vigilance test, the algorithm is able to dynamically create new clusters according to the data distribution when the verification fails, or optimize and update the system when the initial sample recognition is confirmed. The vigilance test consists of calculating a ratio between the distance of the sample to the winning cluster and the average distance of all the samples in this cluster to the cluster centroid,

$$R = \frac{\|\mathbf{x}_j - \mathbf{v}_i\|}{\frac{1}{N_i}\sum_{k=1}^{N_i}\|\mathbf{x}_k - \mathbf{v}_i\|} \tag{6.2}$$

When this ratio is higher than a user-defined threshold, the test fails and a new cluster is created, taking the sample as the initial centroid and assigning an initial cluster distance value to this new cluster (which has only one sample coinciding with the centroid; it is necessary to assign an initial distance value so that the vigilance test can be performed when the next sample is presented). Otherwise, the sample is officially classified into this cluster, and then its centroid and the fuzzy membership values are updated with the following optimization parameters:

$$\mathbf{v}_i = \frac{1}{\sum_{j=1}^{p}(u_{ij})^m}\sum_{j=1}^{p}(u_{ij})^m\mathbf{x}_j \quad i = 1, 2, \ldots, c \tag{6.3}$$

$$u_{ij} = \frac{\left(1/\|\mathbf{x}_j - \mathbf{v}_i\|^2\right)^{1/m-1}}{\sum_{k=1}^{c}\left(1/\|\mathbf{x}_j - \mathbf{v}_k\|^2\right)^{1/m-1}} \quad i = 1, 2, \ldots, c \;\; j = 1, 2, \ldots, p \tag{6.4}$$

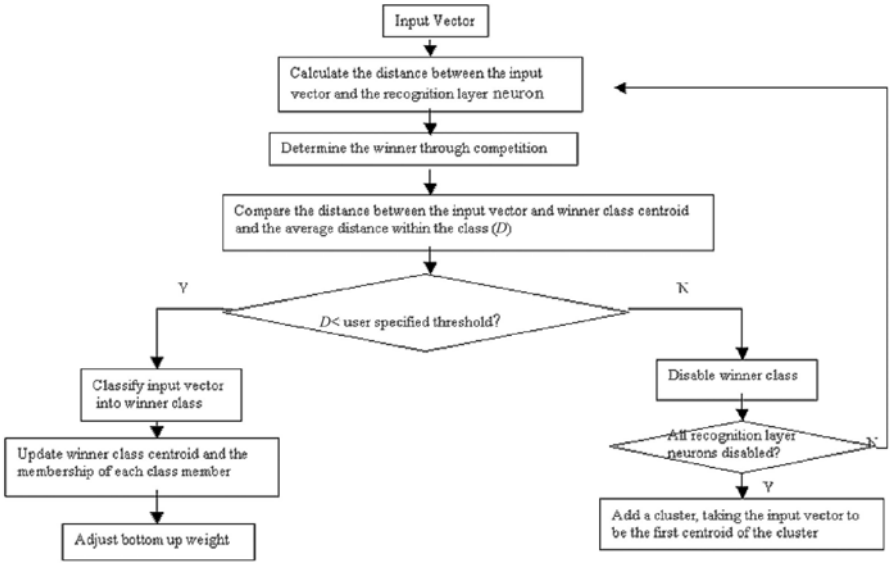


Figure 6.2: Adaptive fuzzy leader clustering (AFLC) implementation flow chart.

The above nonlinear relationships between the ith centroid and the membership value of the jth sample to the ith cluster are obtained by minimizing the fuzzy objective function in Eq. (6.18). Figure 6.2 shows the flow chart for AFLC implementation. AFLC has been successfully applied to image restoration, image noise removal, image segmentation, and compression [13, 31, 32].

6.3.2 Deterministic Annealing

Deterministic annealing [13] is an optimization algorithm based on the principles of information theory, probability theory, and statistical mechanics. Shannon's information theory tells us that the entropy of a system decreases as the underlying probability distribution concentrates and increases as the distribution becomes more uniform. For a physical system with many degrees of freedom that can result in many possible states. A basic rule in statistical mechanics says that when the system is in thermal equilibrium, possibility of a state i follows Gibbs distribution.

$$p_i = \frac{e^{-E_i/k_B T}}{Z} \tag{6.5}$$

where k_B is Boltzmann's constant and Z is a constant independent of all states. Gibbs distribution tells us that states of low energy occur with higher probability than states of high energy, and that as the temperature of the system is lowered, the probability concentrates on a smaller subset of low energy states.

Let E be the average energy of the system, then

$$F = E - TH \tag{6.6}$$

is the "free energy" of the system. We can see that as T approaches zero, F approaches the average energy E. From the principle of minimal free energy, Gibbs distribution collapses on the global minima of E when this happens. SA [33] is an optimization algorithm based on Metropolis algorithm [34] that captures this idea. However, SA moves randomly on the energy surface and converges to a configuration of minimal energy very slowly, if the control parameter T is lowered no faster than logarithmically. DA improves the speed of convergence; the effective energy is deterministically optimized at successively reduced T while maintaining the annealing process aiming at global minimum.

In our clustering problem, we would like to minimize the expected distortion of all the samples x's given a set of centroids y's. Let D be the average distortion,

$$D = \sum_x \sum_y p(x, y) d(x, y) = \sum_x p(x) \sum_y p(y \mid x) d(x, y) \tag{6.7}$$

where $p(x, y)$ is the joint probability distribution of sample x and centroid y, $p(y \mid x)$ is the association probability that relates x to y, and $d(x, y)$ is the distortion measure. Shannon's entropy of the system is given by

$$H(X, Y) = -\sum \sum p(x, y) \log p(x, y) \tag{6.8}$$

If we take D as the average "energy" of the system, then the Lagrangian

$$F = D - TH, \tag{6.9}$$

is equivalent to the free energy of the system. The temperature T here is the Lagrange multiplier, or simply the *pseudotemperature*. Rose [13] described a probabilistic framework for clustering by randomization of the encoding rule, in which each sample is associated with a particular cluster with a certain probability. When F is minimized with respect to the association probability $p(y \mid x)$,

an encoding rule assigning x to y can be obtained,

$$p(y \mid x) = \frac{\exp\left(\frac{-d(x,y)}{T}\right)}{\sum_y \exp\left(\frac{-d(x,y)}{T}\right)} \tag{6.10}$$

Using the explicit expression for $p(y \mid x)$ into the Lagrangian F in Eq. (6.9), we have the new Lagrangian

$$F^* = -T \sum_x p(x) \log \sum_y \exp\left(-\frac{d(x, y)}{T}\right). \tag{6.11}$$

To get centroid values of $\{y\}$, we minimize F^* with respect to y, yielding

$$\sum_x p(x, y) \frac{\mathrm{d}}{\mathrm{d}y} d(x, y) = 0 \tag{6.12}$$

With $p(x, y) = p(x)p(y \mid x)$, where $p(x)$ is given by the source and $p(y \mid x)$ is also known (as given above.) The centroid values of y that minimize F^* can be computed by an iteration that starts at a large value of T, tracking the minimum while decreasing T. The centroid rule is given by

$$y_i = \frac{\sum_x x p(x) p(y_i \mid x)}{p(y_i)} \tag{6.13}$$

where

$$p(y_i \mid x) = \frac{p(y_i)\, e^{-(x-y_i)^2/T}}{\sum_{j=1}^{k} p(y_j)\, e^{-(x-y_j)^2/T}} \tag{6.14}$$

$$p(y_i) = \sum_x p(x) p(y_i \mid x) \tag{6.15}$$

It is obvious that the parameter T controls the entire iterative process of deriving final centroids. As the number of clusters increases, the distortion, or the covariance between samples $\mathbf{x}$ and centroids $\mathbf{y}_i$ will be reduced. Thus, when T is lowered, existing clusters split and the number of clusters will increase while maintaining minimum distortion. When T reaches a value at which the clusters split, it corresponds to a phase transition in the physical system.

The exact value of T, at which a splitting will occur, is given by

$$T_c = 2\lambda_{\max} \tag{6.16}$$

where T_c is known as the critical temperature and $\lambda_{\max}$ is the maximum eigenvalue of the covariance matrix of the posterior distribution $p(x \mid y)$ of the cluster

corresponding to centroid y:

$$C_{x \mid y} = \sum_{x} p(x \mid y)(x - y)(x - y)^t \tag{6.17}$$

In mass-constrained DA, the constraint of $\sum_i p_i = 1$ is applied. Here p_i's are the centroids that coincide in the same cluster i at position y_i. We call this the "repeated" centroids. This is because when the cluster splits, the annealing might result in multiple centroids in each effective cluster depending on the initial pertubation. Below is a simple description of implementation of DA:

1. Initialization: set the maximum number of clusters to be generated $K_{\max}$ and the minimum temperature $T_{\min}$. Set $K = 1$, compute the first T_c and first centroid. Set $py_1 = 1$. Select temperature reduction rate α, perturbation δ, and threshold R.

2. For all current clusters, compute centroids y_i and $p(y_i)$ according to Eqs. (6.13)–(6.15) until converge.

3. Check if $T \geq T_{\min}$, if yes, reduce pseudotemperature $T = \alpha T$, otherwise let $T = 0$ and stop, output centroids and sample assignments.

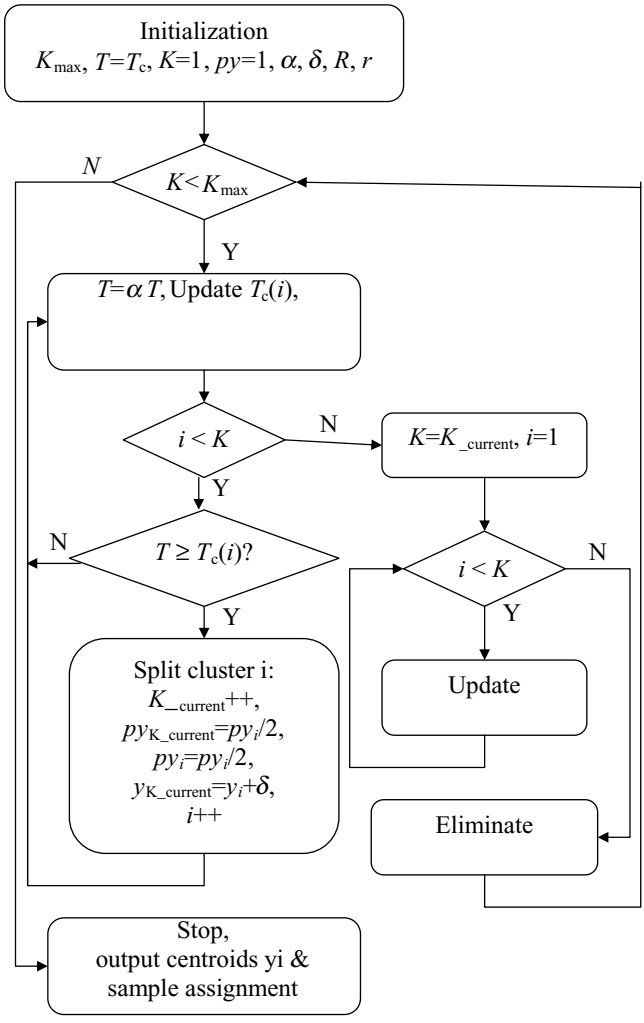
4. If $K > K_{\max}$, stop and output centroids and sample assignments. Otherwise, for all clusters, check if $T > T_c(i)$, if yes, go to step 3; otherwise, split the cluster.

Figure 6.3 gives the flow chart for implementing mass-constrained DA. As can be seen from the flow chart, there are a couple of parameters that govern the annealing process, each exerts its influence on the outcome, particularly the temperature cooling step parameter α. Theoretically, if α is reduced sufficiently slowly, local minima of the cost function can be skipped and a global minimum can be reached. However, it can be very time consuming if T is reduced too slowly.

The centroids y_i and the encoding rule $p(y_i \mid x)$ are illustrated in Fig. 6.4.

6.3.3 k-means Algorithm

k-means clustering [15] partitions a group of samples into K groups. The objective function to be reduced is the sum of square errors. The algorithm iterates as following:



Notes

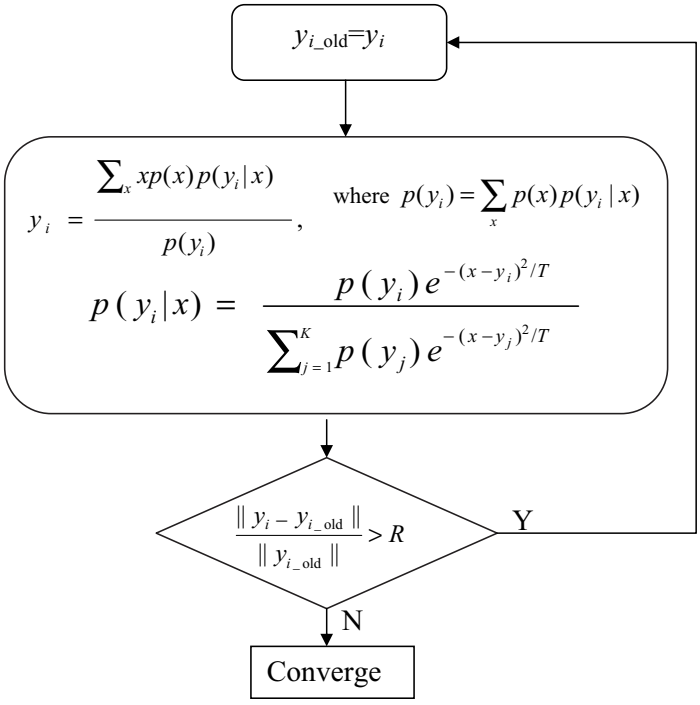
α: temperature cooling rate

δ: perturbation

R: threshold value to eliminate repeated centroids

Update critical temperature: The critical temperature T_c for each cluster at temperature T is calculated as in $T_c=2\lambda max$,

Figure 6.3: Flow chart for implementing mass-constrained deterministic annealing (DA).



Eliminate repeated centroids[*f]: discard centroids that coincide at the same location in one cluster:

For $i=1{:}K$,

If $\|y_i - y_j\|/\|y_i\| < R$, then eliminate y_j, $py_i = py_i + yp_j$, $j=1{:}K$

End

Figure 6.4: Mass-constrained deterministic annealing (DA) centroid update.

1. Compute the mean of each cluster as the centroid of that cluster.
2. Assign each sample to its closest cluster by calculating distances among the sample and all cluster centroids.
3. Keep iterate over the above two steps till the sum of square error of each cluster can no longer be reduced.

The initial centroids can be random; however, the choice of initial centroids is crucial and may result in incorrect partitioning. The iteration drives the objective function toward a minimum. The resultant grouping of the objects is geometrically as compact as possible around the centroids in each group.

6.3.4 Fuzzy C-Means

Fuzzy C-means [16] is a fuzzy version of k-means to include the possibility of having membership of the samples in more than one cluster. The goal is to find an optimal fuzzy c-partition that minimizes the objective function

$$J_m(U, V; X) = \sum_{j=1}^{n}\sum_{i=1}^{c}(u_{ij})^m \|\mathbf{x}_j - \mathbf{v}_i\|^2 \tag{6.18}$$

where $\mathbf{v}_i$ is the centroid of the ith cluster; u_{ij} is the membership value vector of the ith class for the jth sample; d_{ij} is the Euclidean distance between the ith class and sample $\mathbf{x}_j$; c and n denote the number of classes to be clustered and the total number of samples, respectively; and m is a weighting exponential parameter on each fuzzy membership with$1 \leq m < \infty$. The FCM algorithm can be described as follows:

1. Initialize membership function $\mathbf{U}^{(l=0)}$ to random values.
2. Compute the centroid of the ith class with Eq. (6.3).
3. Update membership function $\mathbf{U}^{(l)}$ with Eq. (6.4).
4. If $\left\|\mathbf{U}^{(l-1)} - \mathbf{U}^{(l)}\right\| \leq \varepsilon$ or a predefined number of iteration is reached, stop. Otherwise $l = l + 1$ and go to step 2. ε is a small positive constant.

6.4 Results

We have chosen three different modalities of images, namely, MRI, stereo fundus images, and color cervix images to demonstrate the effectiveness of the advanced clustering algorithms over the traditional ones in segmenting medical images of various modalities.

6.4.1 MRI Segmentation

MRI is one of the most common diagnostic tools in neuroradiology. In brain pathology study, brain and brain tissues have often regions of interest from which abnormality such as the Alzheimer disease or multiple sclerosis (MS) lesions are diagnosed. Numerous techniques in computer-aided extraction of the brain, brain tissues, such as the gray matter, white matter, and cerebrospinal

fluid (CSF), as well as MS lesions have been developed [5, 8, 9, 35–38]. A good survey in applying pattern recognition techniques to MR image segmentation is available in [8]. Clark *et al.* [38] give a comparative study of fuzzy clustering approaches, including FCM and hard c-means versus supervised feedforward back-propagation computational neural network in MRI segmentation. These techniques are found to provide broadly similar results, with fuzzy algorithms showing better segmentation.

MS is a disease that affects the central nervous system. It affects more than 400,000 people in North America. Patients with MS experience range of symptoms depending on where the inflammation and demyelination is situated in the central nervous system. It can be from blurred vision, pain, affecting the sense of touch to loss of muscle strength in arms and legs. About 95% MS lesions occur in the white matter in the brain [39]. MR imaging is usually used to monitor the progression of the disease and the effect of drug therapy. Clinical analysis or grading of MS lesions is mostly performed by experienced raters visually or qualitatively. The involvement of such manual segmentation suffers from inconsistency between raters and inaccuracy. Computer aided automatic or semiautomatic segmentation of MS lesions in MR images is important in enhancing the accuracy of the measurement, facilitating quantitative analysis of the disease [35, 36, 39–43].

Many regular image segmentation techniques can be employed in MS lesion segmentation, such as edge detection, thresholding, region growing, and model-based approaches. However, because of MR field inhomogeneities and partial volume effect, most of the methods are integrated in nature, in which pre- and postprocessing are involved to correct these effects and remove noise, or *a priori* knowledge of the anatomical location of brain tissues is used [36, 39, 41]. Johnston *et al.* [35] used a stochastic-relaxation-based method, a modified iterated conditional modes (ICM) algorithm in 3D [6] on PD- and T2-weighted MR images. Inhomogeneities in multispectral MR images are corrected by applying homomorphic filtering in the preprocessing step. After initial segmentation is obtained, a mask containing only the white matter and the lesion is generated by applying multiple steps of morphological filter and thresholding, on which a second pass of ICM is performed to produce the final segmentation. Zijdenbos *et al.* [36] applied back-propagation neural network for segmentation on both T1-, T2-, and PD-weighted images. Intensity inhomogeneities are corrected by using a so-called thin-plate spline surface fitted to the user-supplied reference points.

Noise is filtered before segmentation is performed by using an anisotropic diffusion smoothing algorithm [44]. An automatic method proposed by Leemput *et al.* [45] removes the need for human interaction by using a probabilistic brain atlas for segmenting MS lesions from T1-, T2-, and PD-weighted images. This method simultaneously estimates the parameters of a stochastic tissue intensity model for normal brain MR images and detects MS lesions as voxels that are not fitted to the model.

6.4.1.1 Normal Brain Segmentation from MRI: Gray Matter, White Matter, or Cerebrospinal Fluid

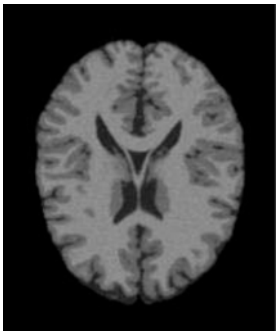
The intensity level and contrast can be very different for T1-, T2-, or PD-weighted MR images. Segmentation of gray matter, white matter, or CSF in the spatial domain depends highly on the contrast of the image intensity; therefore, T1-weighted MRI is more suitable than T2- or PD-weighted MRI. In order to validate the performances of the clustering algorithms, synthetic MRI [46, 47] is used because the existence of an objective truth model is helpful in obtaining quantative analysis of a segmentation technique, excluding the introduction of human error. The synthetic images used in this example are obtained from a simulated brain database [46, 47] provided by McConnell Brain Imaging Center, Montréal Neurological Institute, McGill University. It includes databases for normal brain and MS lesion brain. Three modalities are provided, T1-, T2-, and PD-weighted MRI. Simulations such as noise and intensity nonuniformity are also available.

The image in this example is #90 of 1-mm thick slices with 3% noise and 0% intensity nonuniformity. Figures 6.6–6.9 compare the segmentation results from DA, AFLC, FCM, and k-means. Misclassification, using the computer-generated truth model as the reference, is considered as the performance evaluation criterion following the traditional trend. Misclassification on each segmented category is calculated as the percentage of the total number of misclassified pixels in the segmented image divided by the total number of pixels in the corresponding truth model. For example, for the CSF, let class_csf be the binary segmented image and csf_model be the binary CSF truth model image,

$$\text{N\_miss} = \text{total number of misclassification} = \text{sum (abs (class\_csf-csf\_model))}$$

$$\text{P\_model} = \text{total number of pixels in the CSF truth model} = \text{sum (csf\_model)}$$

$$\text{Misclassification} = \text{N\_miss/P\_model}^{*}100\%$$



(a) T11mm30_90.raw

(b) CSF gray-level truth model

(c) Gray matter gray-level truth model

(d) White matter gray-level truth model

Figure 6.5: Noisy MRI and the corresponding truth model for CSF, gray matter, and white matter.

The truth models are originally fuzzy models (Figs. 6.5(b)–6.5(d)). Since all results produced from the algorithms are hard clustering, the fuzzy truth models are converted into hard models by classifying a pixel to the category in which it has the largest pixel value.

Misclassification results (in Figs. 6.6–6.9) show that DA and AFLC perform better than k-means and FCM, demonstrating the effectiveness of the advanced algorithms in being more noisy tolerant.

6.4.1.2 Segmentation of Lesions in Multiple Sclerosis from MRI

Segmentation of MS lesions from simulated MRI: It is difficult to partition MS from T1-, T2-, or PD-weighted images because of the lack of intensity difference between MS lesion and other brain tissues, as is illustrated in Figs. 6.10(a)–6.10(c). It can be expected that when a pattern recognition technique

(a) Classified CSF

(b) Classified gray matter

(c) Classified white matter

(d) Misclassification: 9.01%

(e) Misclassification: 10.27%

(f) Misclassification: 5.29%

Figure 6.6: Segmentation of noisy MR image by AFLC.

(a) Classified CSF

(b) Classified gray matter

(c) Classified white matter

(d) Misclassification: 10.48%

(e) Misclassification: 10.47%

(f) Misclassification: 5.10%

Figure 6.7: Segmentation of noisy MR image by DA.

(a) Classified CSF (b) Classified gray matter (c) Classified white matter

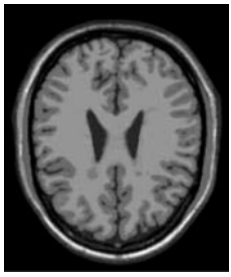
(d) Misclassification: 12.47% (e) Misclassification: 12.08% (f) Misclassification: 5.12%

Figure 6.8: Segmentation of noisy MR image by k-means.

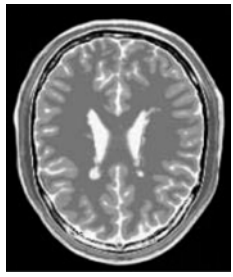
(a) Classified CSF (b) Classified gray matter (c) Classified white matter

(d) Misclassification: 12.36% (e) Misclassification: 12.04% (f) Misclassification: 5.12%

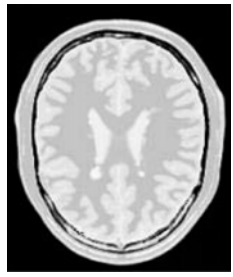
Figure 6.9: Segmentation of noisy MR image by FCM.



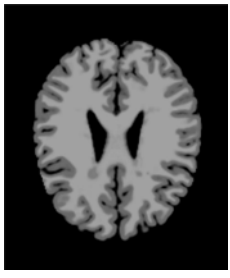
(a) T1-weighted MRI



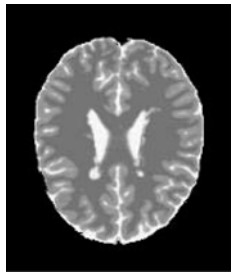
(b) T2-weighted MRI



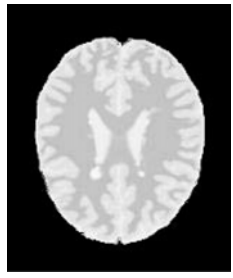
(c) PD-weighted MRI



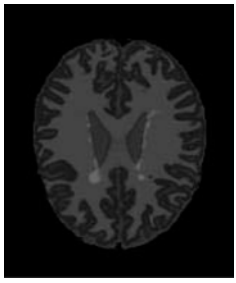
(d) T1-weighted MRI without extraneous parts



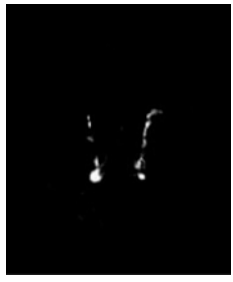
(e) T2-weighted MRI without extraneous parts



(f) PD-weighted MRI without extraneous parts



(g) Synthesized image (T1-(PD-T2))



(h) Fuzzy MS lesion truth model

(i) Hard MS lesion truth model

(j) k-means-segmented MS lesion

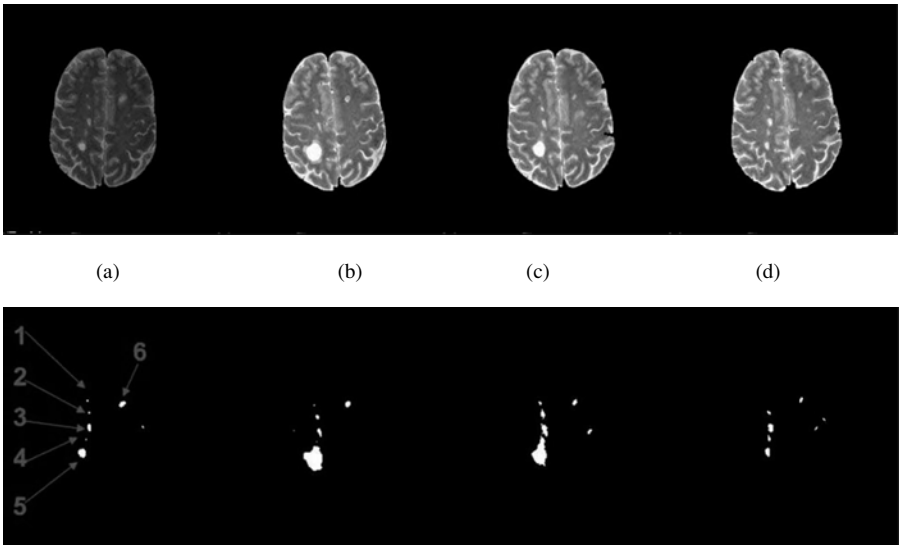
(k) AFLC-segmented MS lesion

(l) DA-segmented MS lesion

Figure 6.10: Segmentation of MS lesions.

is applied, the lesions will be classified either with gray matter in Figs. 6.10(a) and 6.10(c) or CSF in Fig. 6.10(b) since the intensities of the lesions are similar to those tissues. It is a common practice that information embedded in multi-channel MR images are combined to extract MS lesions [35, 36]. In this example, a synthesized image is created by manipulation of the three images: synthesized image = T1 − (PD − T2). Figure 6.10(g) shows the synthesized image of slice #90 T1-, T2-, and PD MR images, respectively. It can be seen that Fig. 6.10(g) provides distinct intensity level variation among gray matter, white matter CSF, and MS lesions. Then synthesized image is feed into four clustering algorithms. Segmentation results are shown in Figs. 6.10(j)–6.10(l). Figure 6.10(h) is the fuzzy MS lesion truth model. A hard model in Fig. 6.10(i) is created by verifying that the fuzzy model possesses the largest value among other tissues at the same pixel. It can be observed that DA provides the closest result to the truth model.

Segmentation of MS lesions from clinical MRI: Clinical MRI is much more complicated than simulated MRI in noise, clarity, and intensity inhomogeneity. The MR images to be segmented in the following example come from clinical data. The T2-weighted MR images are obtained from [33]. Four images are extracted from a MPEG movie showing chronic progression of MS lesions (Figs. 6.11(a)–6.11(d)). The images have been compressed, showing poor image



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 6.11: MR images with MS lesions in chronic order, segmentation and labeling.

resolution and reduced quality (with observable blocking artifacts.) The results of segmentations on these images are summarized in Figs. 6.11(e)–6.11(h), with labeling of the individual lesions shown in Fig. 6.11(e). Segmentation processes will be explained below. With image (a) and (b), intermediate results are shown in Figs. 6.12 and 6.13. For image (c) and (d), intermediate results are skipped and only the white matter masks and final segmentations are illustrated in Figs. 6.14 and 6.15.

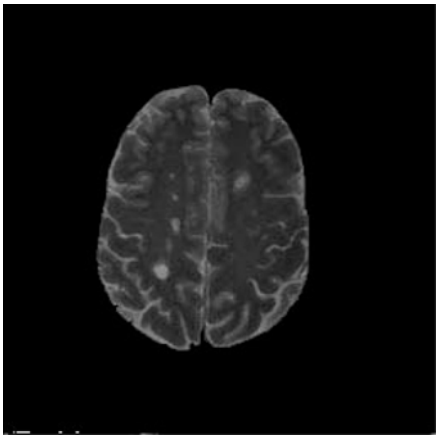
Segmenting MS lesions from a single modality image such as the one used in this example is difficult as has been explained in the previous section. However, when images in other modalities are not available, background knowledge can be used. As most of the MS lesions occur in the white matter, Johnston *et al.* [35] suggested creating a white mask to confine segmentation area such that segmentation accuracy can be enhanced. In this case, segmentation is a two-pass process. The image is first roughly segmented into four categories with DA: the background, the white matter, gray matter, and CSF and other tissues (Figs. 6.12(b)–6.12(e)). Not surprisingly, the MS lesions cannot form a class of their own. They are classified either as gray matter or other categories. Then a white matter mask is generated by morphological filtering of the white matter. This is shown in Fig. 6.12(f). Using this mask, a tailored image containing only the masked area is obtained (Fig. 6.12(g)) and used as input image in the second pass DA clustering. Final segmentation for Fig. 6.12(g) is shown in Fig. 6.12(h). The result is superposed on the original image in Fig. 6.12(i).

Besides being affected by the image quality, the above segmentation is influenced by the mask. The misclassification of gray matter into MS lesion on the right-hand side of the image (indicated by the red arrow in Fig. 6.12) in this example is caused by misclassification of gray matter into the white matter mask.

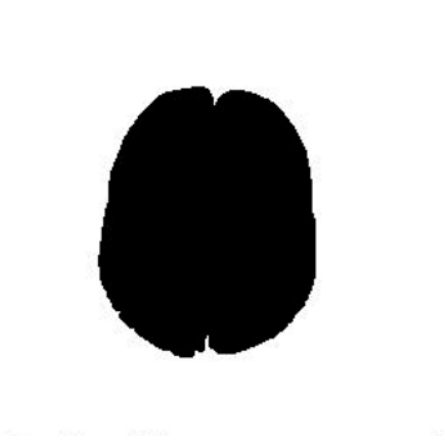
One of the goals to segment the lesions is to provide quantitative analysis. Once the lesions are segmented and labeled, progress of each lesion in size can be obtained in chronic order. Table 6.1 summarizes the changes in size (number of pixels) of all lesions that exist through all four MR images.

6.4.2 Retinal Image Segmentation from Stereo Fundus Images

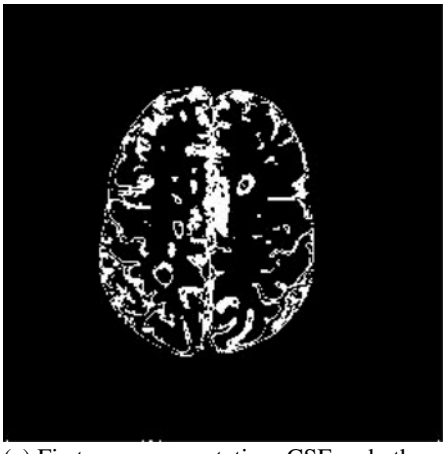
Objects such as blood vessels, optic disk, and optic cup in retinal images are crucial in monitoring and detecting the progression of retinal diseases such as vascular diseases, glaucoma hypertension, and diabetic retinopathy.



(a) Original image

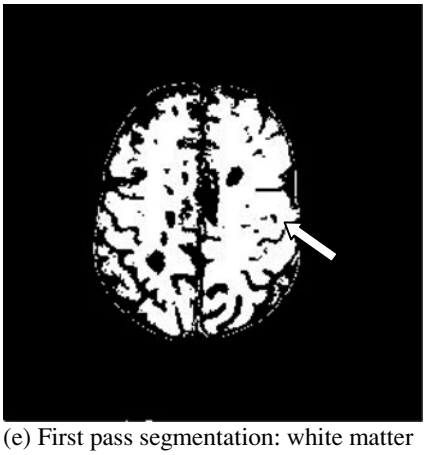


(b) First pass segmentation: background

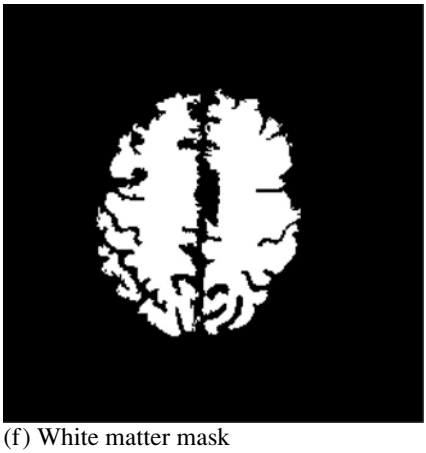


(c) First pass segmentation: CSF and other

(d) First pass segmentation: gray matter

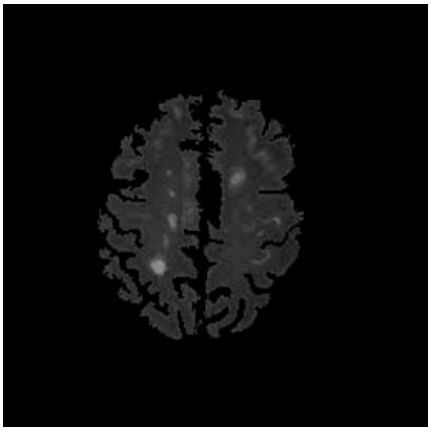


(e) First pass segmentation: white matter

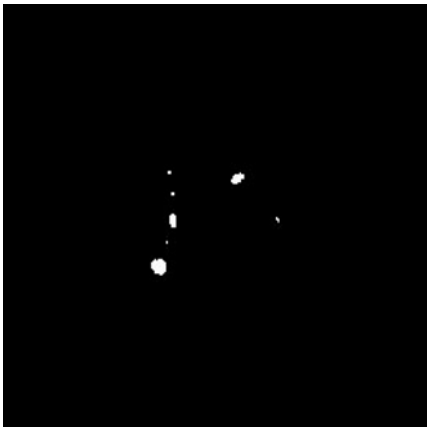


(f) White matter mask

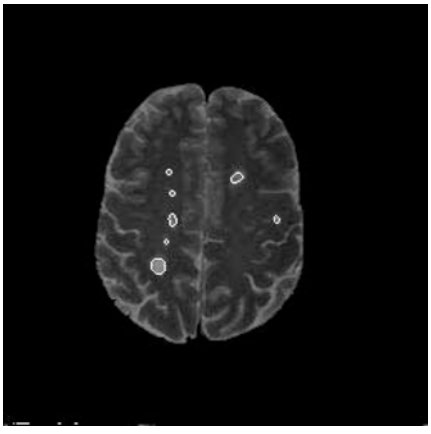
Figure 6.12: DA segmentation of MS lesions from Fig. 6.11(a).



(g) Input image for second pass clustering



(h) MS lesion segmented from the second pass



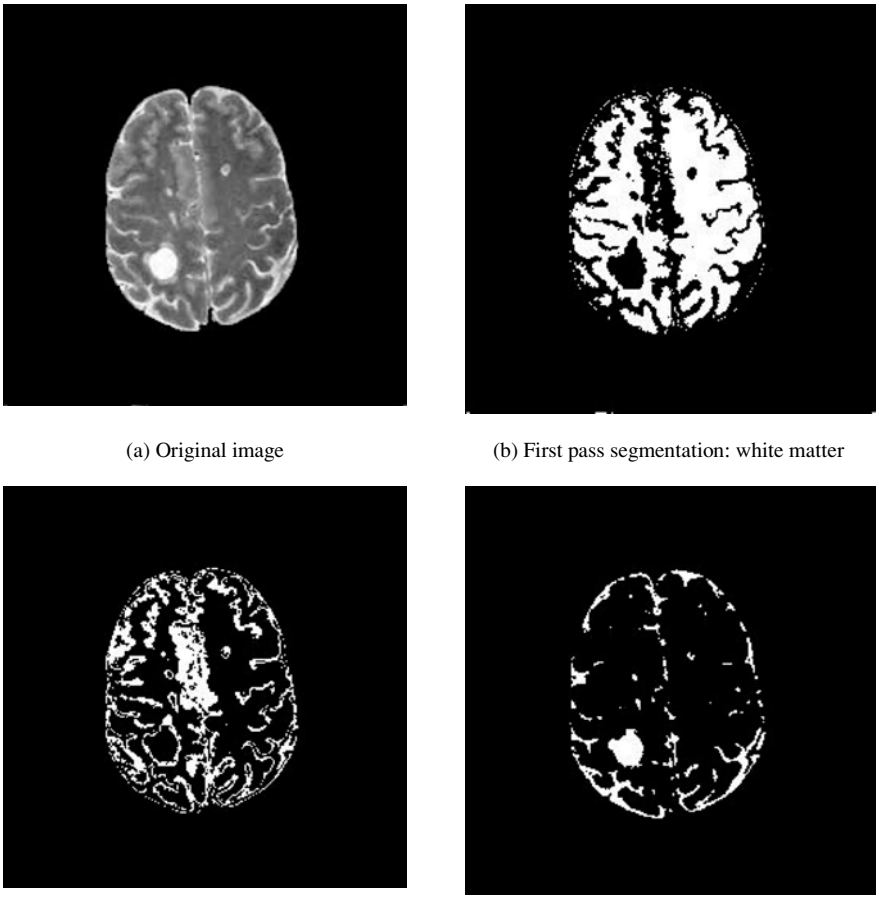
(i) Segmented lesion overlay on (a)

Figure 6.12: (*cont.*)

Table 6.1: Segmented lesion size in chronic MR images

MR image	L1	L2	L3	L4	L5	L6
(a)	4	4	28	2	70	36
(b)	2	15	29	1	401	32
(c)	10	31	63		313[a]	24
(d)	0	11	45	14	41	16

[a] Lesions 4 and 5 merge in (c).



(a) Original image

(b) First pass segmentation: white matter

(c) First pass segmentation: CSF and other

(d) First pass segmentation: gray matter

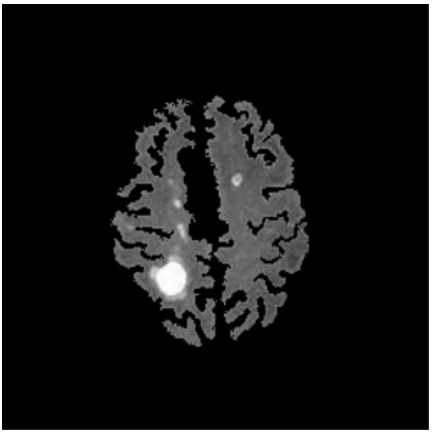
Figure 6.13: DA segmentation of MS lesions from Fig. 6.11(b).

Segmentation of the extracted features allows us to investigate the effect of occlusion induced by these features in generating stereo disparity mapping and 3-D visualization of the optic cup/disk, leading to more accurate diagnosis or monitoring of glaucoma. This is a challenging task due to poor and nonuniform illumination of most fundus images and lack of standardized parameters for stereo imaging geometry.

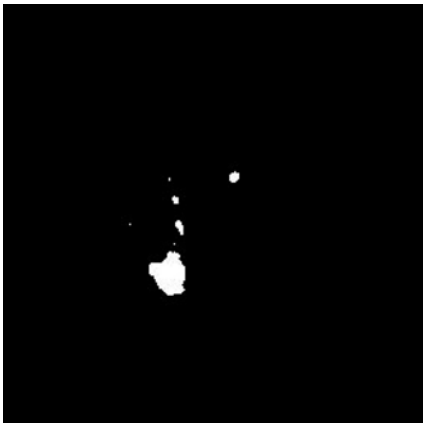
6.4.2.1 3-D Segmentation of the Optic Disk/Cup

The onset and progression of glaucoma can usually be found or monitored through the measurement of changes in the optic disk and optic cup area. It

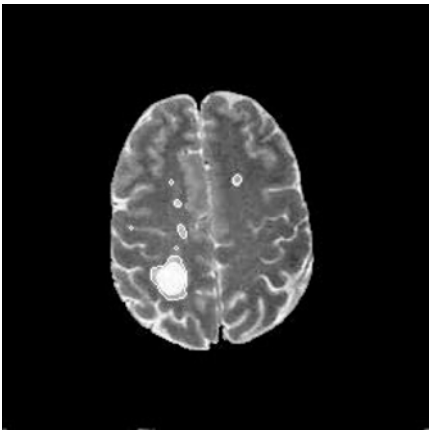
(e) White matter mask



(f) Input image for second pass segmentation



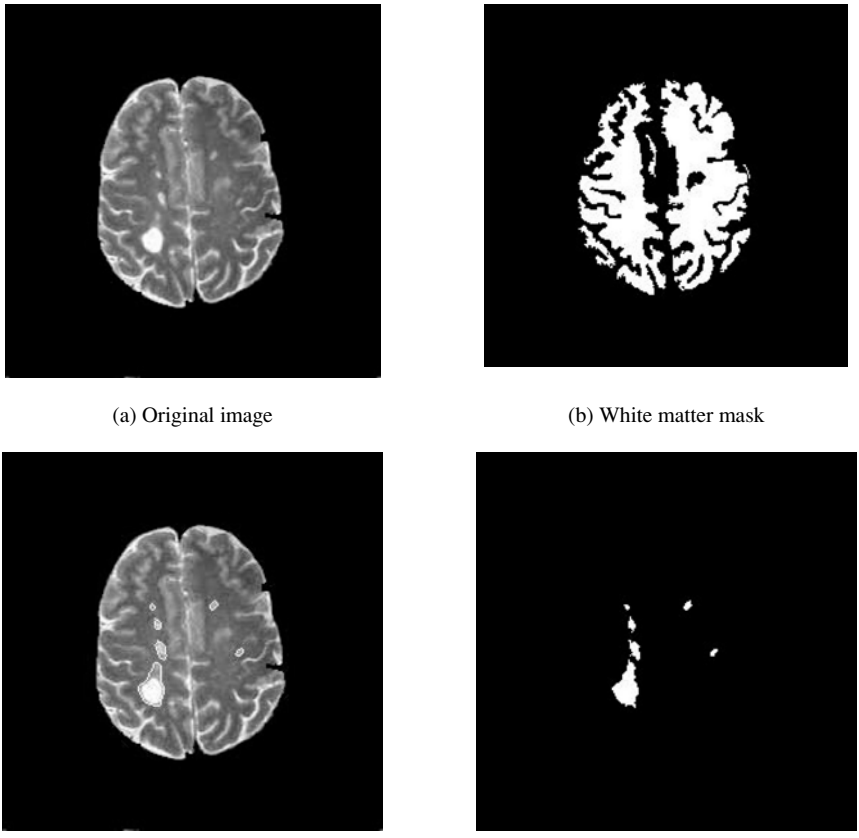
(g) Segmented MS lesions



(h) Segmented lesion overlay on (a)

Figure 6.13: (*cont.*)

can be expressed as the cup-to-disk ratio in diameter (2D) or volume (3D), for which segmentation of the optic cup/disk in 2D or 3D is necessary. The cup-to-disk ratios obtained from 3-D visualization of the optic cup/disk has been found to match closely with those provided by physicians [48]. Semiautomated methods for finding the contours of the optical nerve head (ONH) by digital image analysis attempt to find the disparities of pixels between the fundus stereo pairs in a region including the ONH. Recent studies [48–50] describe in detail the algorithms developed for feature extraction, registration, correlation, and dynamic programming leading to computing disparities based on a nonconvergent stereo



(a) Original image

(b) White matter mask

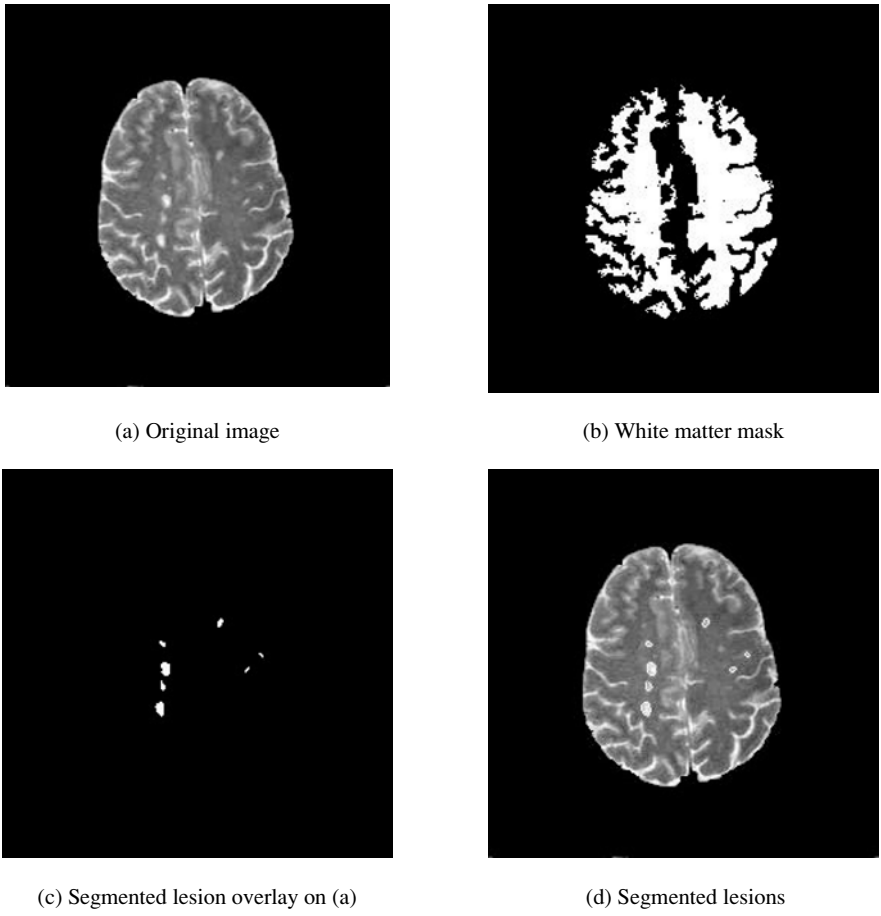
(c) Segmented lesion overlay on (a)

(d) Segmented lesions

Figure 6.14: DA segmentation of MS lesions from Fig. 6.11(c).

imaging system. However, fully automated methods of finding precise disparities between a pair of stereoscopic images may still be problematic due to the presence of noise, occlusions, distortions, and lack of knowledge of the stereo imaging parameters. Under certain constraints, 3-D surface recovery of the optic cup/disk is possible from a pyramidal surface-matching algorithm based on the recovery of the optimum surface within a 3-D cross-correlation coefficient volume via a two-stage dynamic programming technique. The accuracy of the disparity map algorithm leading to 3-D surface recovery is highly dependent on the initial feature extraction process and the stereo imaging parameters.

The stereo images of a real world scene are taken from two different perspectives. The coordinate associated with depth of this scene can be extracted by triangulation of corresponding points in the stereoscopic images [51, 52]



(a) Original image

(b) White matter mask

(c) Segmented lesion overlay on (a)

(d) Segmented lesions

Figure 6.15: DA segmentation of MS lesions from Fig. 6.11(d).

assuming a nonconvergent imaging system. However, the geometrical information of the viewing angles when fundus images are taken in a clinical set-up are not documented, thus introducing certain ambiguities in the disparity mapping of the corresponding points in the stereo pairs. The overall process for 3-D visualization of retinal structures from stereo image pairs is complex and includes different matching strategies, area or feature based or a combination of both. Several preprocessing steps are also followed for feature extraction and registration prior to coarse to fine disparity search.

In the preprocessing stage, three channel (RGB) decomposition is first performed on the original color pair. Only the green channel is processed since it is the one that carries the most information. Red and blue channels have low

entropy in relation to the green channel and therefore are not taken into account. The registration process removes all vertical displacements leaving only the horizontal shifts arising from the different positions of the camera while taking the stereo fundus images. A good registration is crucial to obtain accurate disparity maps. A power cepstrum-based registration that uses Fourier spectrum properties to correct rotational errors that may be present in the stereo pair is employed. This process begins by extracting the most relevant features such as the blood vessels in both images. These features are extracted by subtracting a filtered version of the original stereo pair from the original (unsharp masking). After binarizing this new stereo pair, multiple passes of a median filter are used to eliminate some of the resulting noise in the images. Compensation for rotational differences is also performed via zero mean normalized cross-correlation(ZNCC) [53] of the Fourier spectrum of the images using ZNCC as a disparity measure. ZNCC is expressed as follows:

$$C(i, j) = \frac{\mathrm{cov}_{i,j}(f, g)}{\sigma_{i,j}(f) \times \sigma_{i,j}(g)} \tag{6.19}$$

$$\mathrm{cov}_{i,j}(f, g) = \frac{1}{((2K+1)(2L+1)-1)} \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} \left(f_{m,n} - \overline{f_{i,j}}\right)\left(g_{m,n} - \overline{g_{i,j}}\right) \tag{6.20}$$

where f and g are the windows of pixels to be measured and $\overline{f}$ and $\overline{g}$ are corresponding average values. K and L define the size of those windows, and the indices for the pixels within the windows are i and j. $\sigma(f)$ and $\sigma(g)$ are the square roots of the covariances $\mathrm{cov}(f, f)$ and $\mathrm{cov}(g, g)$, respectively.

According to the inherent Fourier spectrum properties, a rotation in the spatial image results in the same amount of rotation of its spectrum. Thus it is possible to find the angle of rotation of one image in the stereo pair with respect to the other by performing step-by-step rotations and cross correlating their Fourier transforms. The actual angle of rotation will be the one with the highest cross-correlation obtained. Rotational compensation is applied once the angle of rotation has been found.

After the rotational correction, a cepstrum transformation is applied to the sum of the binary-featured stereo pair images. The power cepstrum P is defined as in [52]:

$$P\left[i(x, y)\right] = |\mathbf{F}(\ln\{|\mathbf{F}[i(x, y)]|^2\})|^2 \tag{6.21}$$

where **F** represents the Fourier transform operation. Let $w(x, y)$ be the reference image, $w(x + x_0, y + y_0)$ be the shifted image, and $i(x, y) = w(x, y) + w(x + x_0, y + y_0)$. Then the power cepstrum of the sum of both images is given as

$$P[i(x, y)] = P[w(x, y)] + A\delta(x, y) + B\delta(x \pm x_0, y \pm y_0) + C\delta(x \pm 2x_0, y \pm 2y_0) + \cdots \quad (6.22)$$

where $\delta(x, y)$ is the Kronecker delta and A, B, and C are the first three coefficients for this power cepstrum expansion series [54]. Equation (6.22) shows that the displacement between images results in the sum of the power cepstrum of the original image $w(x, y)$ plus a multitude of delta functions. Each delta is separated from the others by an integer multiple of the actual displacement we are looking for. In order to enhance the cepstral peaks, the cepstrum of the reference must be subtracted from the cepstrum of the stereo pair. With this in mind, a fixed number of deltas are chosen from the resulting cepstrum. Each delta represents a translational shift, or an integer multiple of the shift, of a pixel in the image shifted from the corresponding pixel in the reference image. All points are tested by cross correlating the reference image with the other image shifted by the number of pixels (in the vertical and horizontal directions) indicated by the current point being tested. The highest correlation will correspond to the most probable relative translation between both images.

Before disparity mapping is carried out, salient features of the image, such as the blood vessels, have to be extracted. The blood vessels are segmented through series of traditional local operation, such as unsharp masking, thresholding, and median filtering. This process is illustrated in Fig. 6.16.

The algorithm developed for the search of disparities first divides both images into square windows of a given size (multiple of two), say N by N. ZNCC is performed between the windows in one image with the windows in the other image. If cross correlation is larger than a certain threshold, it is assumed that the windows at that position in the image are similar, so the cepstrum is applied to those windows to check for possible shifts. Otherwise, if the cross correlation is smaller than or equal to the threshold mentioned, zero disparity is assigned to every pixel in the window. Only a specified number of horizontal points shown in the cepstrum are taken into account for analysis. Let's say that, for an N by N window, only $N/4$ horizontal points are chosen for analysis in the cepstral plane. This is because for an N by N window the maximum horizontal displacement that can be detected is $N/2$ (either to the right or to the left, making a full range

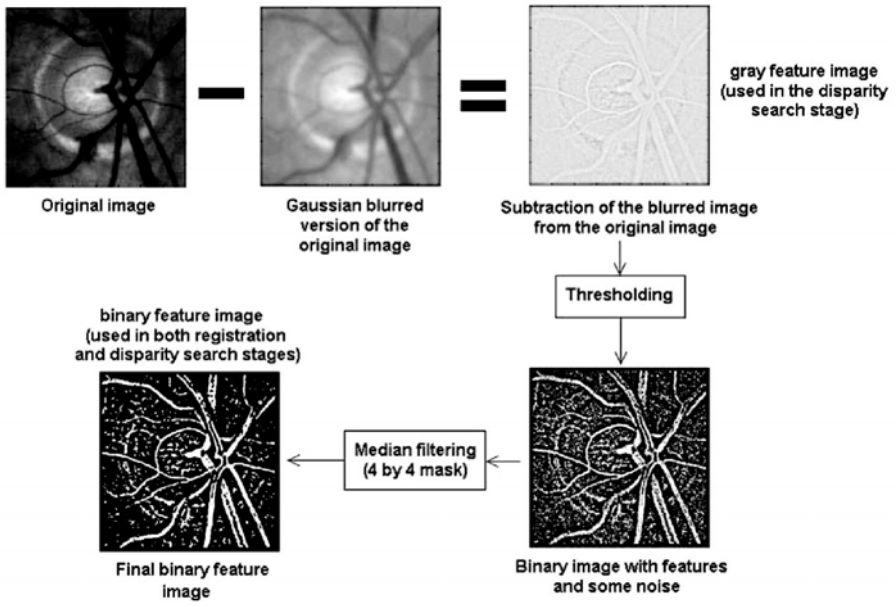
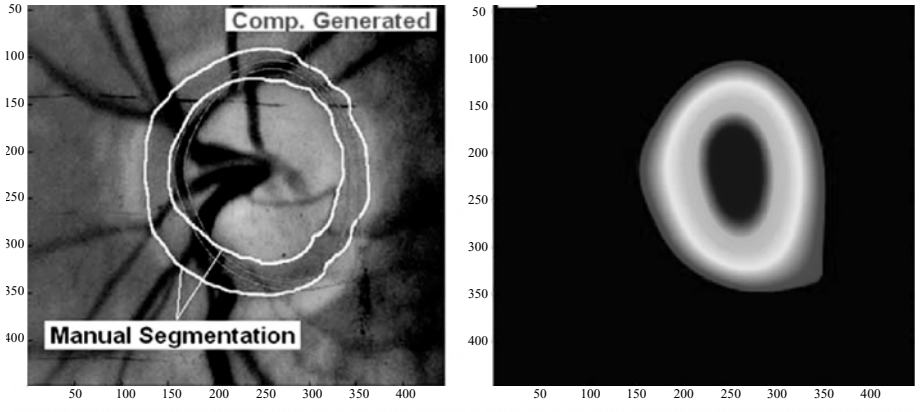


Figure 6.16: Blood vessel extraction for registration and disparity mapping.

of N pixels), so checking all $N/2$ points for right and left shifts will be very time consuming. Instead, only the most probable $N/4$ horizontal shifts found by the cepstrum will be tested using the cross correlation technique. One of the images of the stereo pair is considered as the reference image and the other is the test image. Then, for every point chosen (from the cepstrum), cross correlation is applied between the reference window (in the reference image) and the other window (in the test image) shifted by the number of pixels determined by the cepstral shift. Since the cepstrum can detect only the amount of the shifts but not their direction (the cepstrum is symmetric about the origin), each point should be tested for left and right shifts. So when checking $N/4$ cepstral points, actually $N/2$ positions are analyzed. The highest value in the cross correlation will be the most probable shift that will be assigned to all elements in the window currently being tested for disparity. The number of cepstral points is not a fixed parameter and can be modified. This modification will affect the processing time and the accuracy of the disparity map. Once all disparities have been calculated with a window size of N by N pixels, the size of the window is reduced by a factor of two and the whole process is repeated until the windows reach a predetermined size. Each disparity map (calculated at a given resolution) is accumulated by adding it to the previous disparity map. At the end of the process, the final disparity map

is the total accumulated disparity map. Usually the starting window size is 64 by 64 and the stopping size is 8 by 8. Smaller sizes of a window may not be worth computing because of the much longer time required and the small impact of it on the final disparity map. Also, since the window is so small, chances are that noise becomes a serious issue. The cepstrum is, in fact, a very noise tolerant technique that is suitable for finding disparities in chosen regions [55], while cross correlation is noise-sensitive and finds disparities using a procedure in a pixel-by-pixel fashion. A combination of both techniques results in an accurate and noise tolerant algorithm. In order to get an accurate 3-D representation from a stereo pair of images, disparities must be known for each point (pixel) of one image with respect to the other. Since the disparity search algorithm finds only disparities for the features or regions, disparities of all individual pixels are not known. The interpolation used here gives an estimate of the other missing disparities. Cubic B-spline is the interpolation technique applied to the sparse matrix resulting from the disparity search. It can be shown that the cubic B-spline can be modeled by three successive convolutions with a constant mask [54, 56]. In this case, a mask consisting of all ones of size 32 by 32 or 64 by 64 is used. After filtering the original sparse disparity matrix three times with the mask described above, a smooth representation results. This is the final 3-D surface of the ONH. With this surface, measures such as the disk and cup volume can be made. Figure 6.17(a) shows the optic disk/cup segmentation obtained from the



(a) (b)

Figure 6.17: Segmentation of optic disk/cup from 3-D disparity map. (color slide)

above disparity mapping. The manual segmentation from an ophthalmologist is also shown for reference. Figure 6.17(b) is the smoothed disparity maps from which the iso-disparity contours were obtained.

6.4.2.2 Blood Vessel Segmentation via Clustering

Blood vessel segmentation has been a well-researched area in recent years, motivated by the needs such as image registration in our case, as well as in visualization and computer-aided surgery. A large number of vessel segmentation algorithms and techniques have been developed, oriented toward various medical image modalities, including X-ray angiography, MRI, computed tomography (CT), and other clinically used images. Like general image segmentation, vessel segmentation is also an application and image modality specific. It can be automatic or nonautomatic. According to a very thorough recent survey [57], vessel segmentation techniques can be roughly classified into the following categories (again, similar to general image segmentation, some techniques are integrated methods that combine approaches from different categories):

1. Pattern recognition techniques: This category includes multiscale approaches that extract large and fine vessels through low resolution and high resolution, respectively; skeleton-based approaches that extract blood vessel centerlines and connect them to form a vessel tree, which can be achieved through thresholding and morphological thinning; region-growing approaches that group nearby pixels that are sharing the same intensity characteristics assuming that adjacent pixels that have similar intensity values are likely to belong to the same objects; ridge-based approaches that map 2-D image into a 3-D surface, and then detect the ridges (local maximums) using various methods; differential geometry-based approaches; matching filter approaches; and morphological approaches.

2. Model-based methods: Model-based methods extract vessels by using explicit vessel models. Active contour (or snake) [58] finds vessel contours by using parametric curves that changes shapes when internal and external forces are applied. Level set theory [59] can also be adapted to vessel segmentation. Parametric models, template matching, and generalized cylinders model also fall into this category.

3. Tracking-based methods: Tracking-based approaches track a vessel from a starting point on a vessel, and detect vessel centerlines or boundaries by analyzing the pixels orthogonal to the tracking direction.

4. Artificial intelligence methods: These methods use the knowledge of targets to be segmented, for example, the properties of the image modality, the appearance of the blood vessels, anatomical knowledge, and other high-level knowledge as guidelines during the segmentation process. Low-level image processing techniques such as thresholding, simple morphological operations, and linking are employed depending on the guidelines.

5. Neural network-based methods: Neural networks are designed to learn the features of the vessels via training, in which a set of sample vessels containing various targeted features are fed into the network such that the network can be taught to recognize the object with these features. The capability of the trained network in recognizing vessels depends on how thorough the features are represented in the training set and how well the network is adjusted.

6. Tube-like object detection methods: This category represents other methods that extract tube-like objects and can be applied to vessel segmentation.

Some of the above approaches have been applied to retinal blood vessel segmentation. Chaudhuri *et al.* [60] used a tortuosity measurement technique and matching filter for blood vessel extraction and reported 91% of blood vessel segments and 95% of vessel network. Wood *et al.* [61] extracted blood vessels in the retinal image for registration by first equalizing the image with local averaging and subtraction, and then nonlinear morphological filtering was used to locate blood vessel segments. Hoover *et al.* [62] proposed an automated method to locate blood vessels in images of the ocular fundus. They used both local and global vessel features and studied the matched filter response using a probing technique. Zana and Klein [63] segmented vessels in retinal angiography images based on mathematical morphology and linear processing. FCM and tracking-based methods can also be used in retinal vessel segmentation [64]. Zhou *et al.* [65] proposed an algorithm that relies on a matched filtering approach coupled with a priori knowledge about retinal vessel properties to automatically detect

the vessel boundaries, track the midline of the vessel, and extract useful parameters of clinical interest.

We can see from the previous section that accurate blood vessel segmentation is very important for registration and disparity mapping. The segmentation method described previously involves Gaussian filtering, unsharp masking, thresholding, and median filtering all of which can be easily affected by image illumination or contrast of the image. Here, we used advanced clustering approaches for the same task and compared it with the Gaussian filtering method.

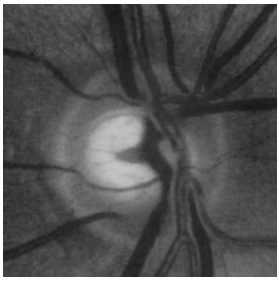
Among the three color channels, the green channel provides better contrast for edge information. Therefore, it alone can be used to accelerate subsequent segmentation processes. Challenge of blood vessel segmentation lies in distinguishing the blood vessel edges. However, most images are noisy and nonuniformly illuminated. In the optic disk area, blood vessel edges are smeared by reflectance. Simple segmentation techniques such as the one described above can produce noisy and inaccurate result. As is illustrated in Fig. 6.19(a), edges of the optic disk are mistakenly classified as blood vessel edges.

DA clustering solves this problem. Figure 6.18(c) shows the segmented result. Image resulted directly from DA segmentation is still affected by the noise in the original image, since single pixel intensity is used as feature. However, the noise in the segmented image can be easily removed through morphological filtering. The expansion or shrinking of blood vessel edges caused by morphological erosion and dilation can be corrected by edges detected by a simple edge detector, such as a Canny edge detector. Figure 6.18(d)–6.18(f) show the segmented optic disk, optic cup, and blood vessels in the ONH, respectively. The optic disk/cup segmentation is very similar to the manual segmentation in Fig. 6.18(b).

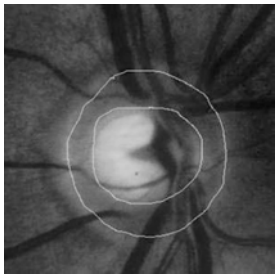
Blood vessels thus segmented are then used for registration and disparity mapping in the 3-D optic disk/cup segmentation, as is shown in Fig. 6.19(b). Three-dimensional visualization of the optic disk/cup with and without DA segmentation is comparatively demonstrated in Figs. 6.19(c) and 6.19(d). We can see that with DA segmentation, more accuracy is achieved.

6.4.3 Color Cervix Image Segmentation

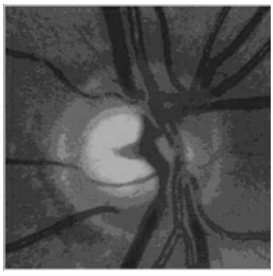
Cervical cancer is the second most common cancer among women worldwide. In developing countries, cervical cancer is the leading cause of death from cancer. About 370,000 new cases of cervical cancer occur worldwide, resulting



(a) Fundus image (left stereo image)



(b) Manually segmented optic disk/cup by an ophthalmologist on the right stereo image

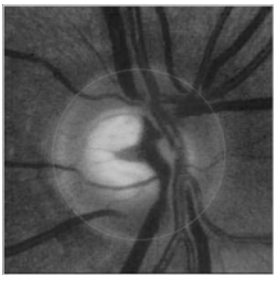


(c) DA segmentation of (b)

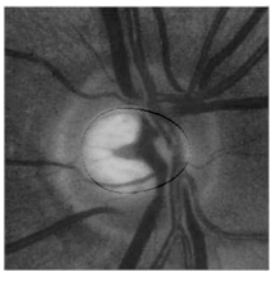
(d) DA-segmented optic disk

(e) DA-segmented optic cup

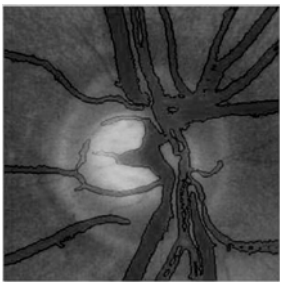
(f) DA-segmented blood vessels



(g) Final segmentation of optic disk



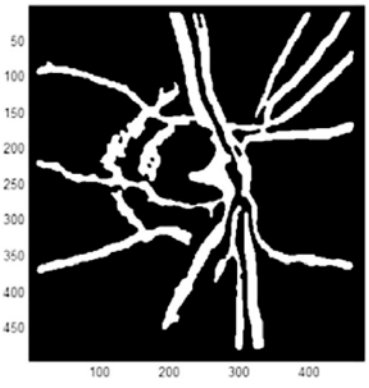
(h) Final segmentation of optic cup



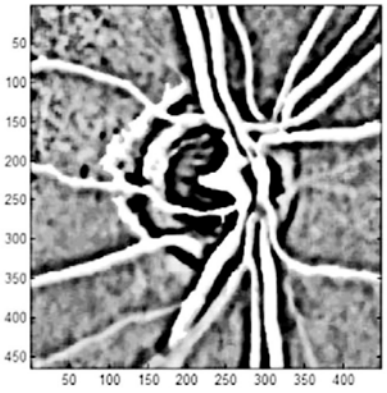
(i) Final segmentation of blood vessels

Figure 6.18: DA segmentation on clinical retinal image.

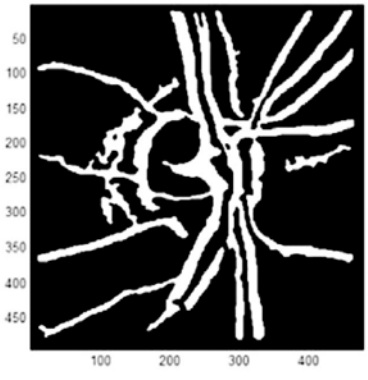
around 230,000 deaths each year. Cervical cancer develops slowly and has a detectable and treatable precursor condition known as dysplasia. It can be prevented through screening at-risk women and treating women with precancerous and cancerous lesions. In many western countries, cervical cancer screening programs have reduced cervical cancer incidence and mortality by as much as 90%. Analysis and interpretation of cervix images are important in early detection



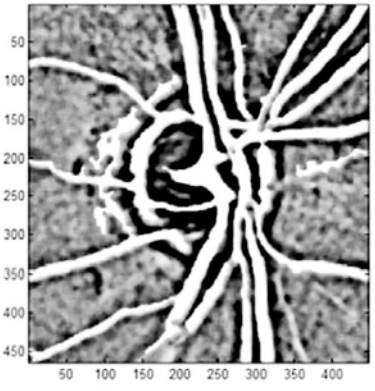
Segmented blood vessels (left)



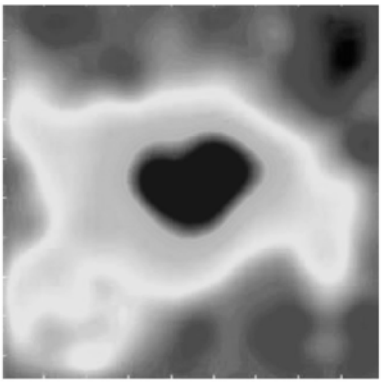
Feature map (left)



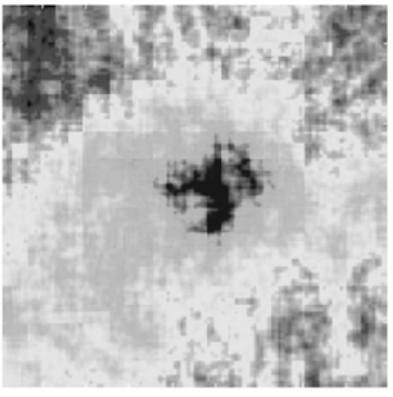
Segmented blood vessels (right)



Feature map (right)



Interpolated disparity map

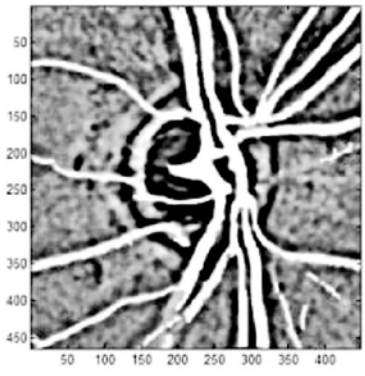


Disparity map

(a) Disparity map with old segmentation technique

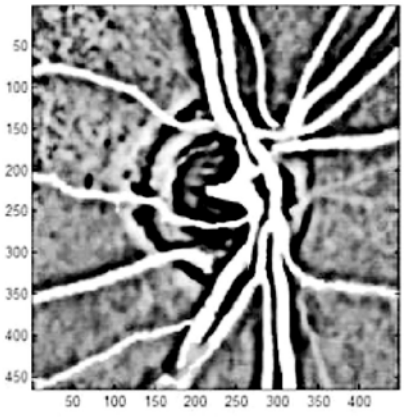
Figure 6.19: Disparity maps generated with and without DA feature extraction.

Segmented blood vessels (left)

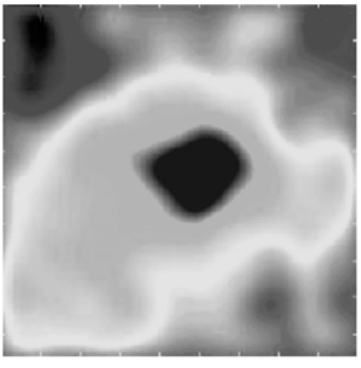


Feature map (left)

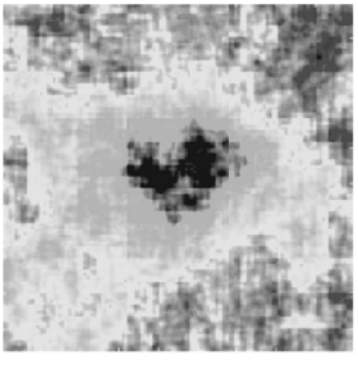
Segmented blood vessels (right)



Feature map (right)

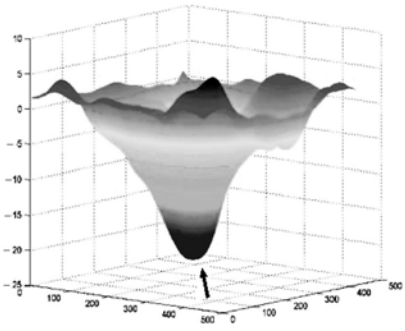


Interpolated disparity map

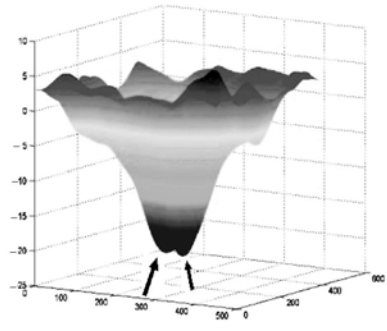


Disparity map

(b) New DA-based segmentation technique



(c) Disparity map obtained from DA blood vessel segmentation



(d) Disparity map obtained from general edge detection for blood vessel segmentation

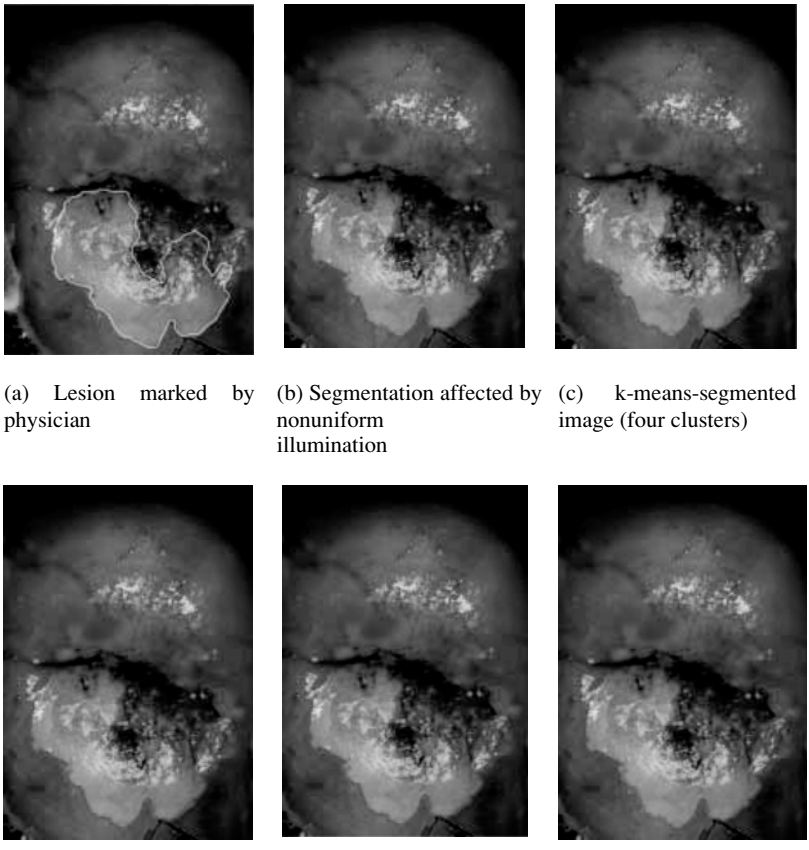
Figure 6.19: (*cont.*)

of cervical lesions. Automated image analysis is helpful in providing quantative lesion description thus monitoring of chronic lesions so that the onset of cervical cancer can be treated effectively.

A cervix image is a magnified color photograph of the cervix (illustrated in Fig. 6.20(a)). The acetowhite lesion area below the opening is marked by a trained physician, serving as a reference to other segmented images resulting from the algorithms. This image is taken with a regular high-resolution color camera, thus the most prominent problems preceding segmentation of the acetowhite lesion area from the rest are the reduction or removal of the glare and non-uniform illumination. Figure 6.20(b) shows the segmentation without illumination correction. All algorithms fail to recognize the lesion close to the cervix opening because the area is darker than other parts of the lesion, and vice versa for the section in the lower part of the cervigram, where the normal area is falsely classified as lesion. The glare on the top left of the lesion is also misclassified. Figures 6.20(c)–6.20(f) are the segmentation results from k-means, DA, and AFLC after illumination correction. The results are similar, with DA generating the closest partition to the manual segmentation.

6.5 Conclusions

From segmentation of the MR images, it can be observed that DA provides the best performance in terms of accuracy and stability among all discussed



(a) Lesion marked by physician

(b) Segmentation affected by nonuniform illumination

(c) k-means-segmented image (four clusters)

(d) DA-segmented lesion (four clusters)

(e) AFLC-segmented lesion (eleven clusters)

(f) AFLC-segmented lesion (four clusters)

Figure 6.20: Segmentation of cervical lesion.

clustering algorithms in several aspects. It is unsupervised; since theoretically it is designed to reach global minimum, the result is not biased by initialization. For the same type of image, the pseudotemperature reduction rate can be fixed, thus the segmentation process does not need parameter manipulation thereby yielding fully automated processing. Although the processing speed of DA is slower than other clustering algorithms, for small images such as the ones used above (217×181), the processing speed of DA is comparable to the other algorithms used. DA is also noise tolerant because of its statistical nature.

Both k-means and FCM, the well-known clustering algorithms, suffer from the initialization and local minimum problems. Cluster initialization is crucial in

yielding satisfactory results. When not initialized properly, a clustering algorithm might be trapped in a local minimum, failing to proceed to the correct cluster. Our experimentations show that with random initialization, both k-means and FCM fail to generate the lesion clusters in MRI MS segmentation. AFLC is an automated and adaptive improvement over k-means and FCM by incorporating neural leader clustering and FCM. The performance is improved; however, similar problems are still encountered. Initialization is eliminated by selecting the first incoming sample as initial centroid, therefore, the outcome is sample-order dependent. DA is the best candidate for medical image segmentation by an advanced clustering technique. It is not sensitive to parameter tuning, and initialization problem, and is noise tolerant and guaranteed to converge.

Advanced clustering techniques can provide general solutions for effective segmentation of a broad range of medical images. All segmentation examples presented in section 6.3 use image intensity as the single feature to clustering algorithms to demonstrate the efficiency of the algorithms. In real applications, local property or connectivity of adjacent pixel can be embedded into segmentation to achieve more accurate segmentation [66, 67].

6.6 Acknowledgments

This work has been partially supported by funds from the Advanced Technology Program (ATP) (Grant No. 003644-0280-1999), Technology Development and Transfer Program (TDT) (Grant No. 003644-0217-2001) of the state of Texas, Kestrel Corporation, the NEI Grant No. 1 R43 EY14090-01 and the NSF Grant EIA-9980296. We acknowledge Young I. Kim, M.D., and Mary Lucy M. Pereira, M.D., of Young H. Kwon's (M.D., Ph.D.) team from University of Iowa Hospitals and Clinics for their contributions to manual segmentation of the stereo optic disk images. The authors are grateful to Daron Ferris, M.D., from the Medical College of Georgia for providing us with the cervigram images from the Guanacaste Project, Costa Rica, sponsored by the National Cancer Institute of USA.

Questions

1. *What is the structure of adaptive fuzzy leader clustering (AFLC)?*
2. *Does AFLC have to initialize like k-means? If not, why?*

3. *How does AFLC dynamically adjust the number of clusters?*
4. *What is the difference between deterministic annealing (DA) and simulated annealing (SA)?*
5. *What is the DA cost function and what does it minimize?*
6. *What effect does the temperature reduction rate parameter have on DA clustering?*
7. *How does DA adjust the number of clusters?*
8. *What does mass-constrained DA mean?*
9. *What makes MS segmentation different from normal brain segmentation?*
10. *Judging from the examples given in the chapter, what are the performance differences among AFLC, DA, FCM, and k-means?*
11. *What is the limitation of clustering segmentation based on image intensity?*
12. *How is clustering in retinal optic disk/cup and blood vessel segmentation better than regular edge detection techniques?*
13. *Why is registration necessary in 3-D retinal disk/cup segmentation and how is it done?*
14. *How is the 3-D optic disk/cup map created?*

Bibliography

[1] Pham, D. L., Xu, C. J., and Prince, L., A survey of current methods in medical image segmentation, Ann. Rev. Biomed. Eng., Jan 1998.

[2] Haralick, R. M. and Shapiro, L. G., Image segmentation techniques, Comput. Vision, Graphics Image Process., Vol. 29, No. 1, pp. 100–132, 1985.

[3] Fu, K. S. and Mui, J. K. A survey on image segmentation, Patt. Recogn., Vol. 13, pp. 3–16, 1981.

[4] Pal, N. R. and Pal, S. K., A review on image segmentation techniques, Patt. Recogn., Vol. 26, No. 9, pp. 1277–1294, 1993.

[5] Suri, J. S., Setarehdan, S. K., and Singh, S., eds., Advanced Algorithmic Approaches to Medical Image Segmentation, Springer-Verlag, London, 2002.

[6] Skarbek, W. and Koschan, A., Colour image segmentation: A survey, Technical Report 94-32, Technical University, Berlin, 1994.

[7] Lucchese, L. and Mitra, S. K., Color image segmentation: A state-of-the-art survey, Proc. Indian Nat. Sci. Acad. (INSA-A), Vol. 67, No. 2, pp. 207–221, 2001.

[8] Bezdek, J. C., Hall, L. O., and Clarke, L. P., Review of MR image segmentation techniques using pattern recognition, Med. Phys., Vol. 20, pp. 1033–1048, 1993.

[9] Clarke, L. P., Camacho, R. P., Velthuizen, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., and Silbiger, M. L., Review of MRI segmentation: Methods and applications, Magn. Reso. Imaging, Vol. 13, No. 3, pp. 343–368, 1995.

[10] Zadeh, L. A., Fuzzy sets, Inf. Control Theory, Vol. 8, pp. 338–353, 1965.

[11] Zadeh, L. A., Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst., Man, Cybern., Vol. SMC-3, pp. 28–44, 1973.

[12] Jain, A. K., Murty, M. N., and Flynn, P. J. Data clustering: A review, ACM Comput Surveys, Vol. 31, No. 3, pp. 264–323, 1999.

[13] Rose, K. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems, Proc. IEEE, Vol. 86, No. 11, 1998.

[14] Newton, S. C., Pemmaraju, S., and Mitra, S., Adaptive fuzzy leader clustering of complex data sets in pattern recognition, IEEE Trans. Neural Networks, Vol. 3, pp. 794–800, 1992.

[15] MacQueen, J., Some methods of classification and analysis of multivariate observations, In: Proceedings of 5th Berkeley Symposium on Math., Stat., and Prob., LeCam, L. M. and Neyman, J., eds., University of California Press, Berkeley, CA, pp. 281, 1967.

[16] Bezdek, J., Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

[17] Gonzalez, R. C. and Woods, R. E., Digital Image Processing, Addison Wesley, Reading, MA, 1992.

[18] Marroquin, J. L. and Girosi, F., Some extensions of the k-means algorithm for image segmentation and pattern recognition, Technical Report, MIT AI Lab., AI Memo 1390, Jan 1993.

[19] Fraley, C. and Raftery, A. E., How many clusters? Which clustering method? Answers via model-based cluster analysis, Technical Report No. 329, University of Washington, 1998.

[20] Ray, S., Turi, R. H., and Tischer, P. E., Clustering-based color image segmentation: An evaluation study, In: Proceedings of Digital Image Computing: Techniques and Applications, Brishbane, Qld., Austrlia, 6–8 Dec. 1995, pp. 86–92.

[21] Park, S. H., Yun, I. D., and Lee, S. U., Color image segmentation based on 3-D clustering: Morphological approach, Patt. Recogn., Vol. 31, No. 8, pp. 1061–1076, 1998.

[22] Weeks, A. R., and Hague, G. E., Color segmentation in the HIS color space using the k-means algorithm, In: Proceedings of the

SPIE—Nonlinear Image Processing VIII, San Jose, CA, Feb, 10–11, 1997, pp. 143–154.

[23] Wu, J., Yan, H., and Chalmers, A. N., Color image segmentation using fuzzy clustering and supervised learning, J. Electron. Imaging, pp. 397–403, 1994.

[24] Pappas, T. N., An adaptive clustering algorithm for image segmentation, IEEE Trans. Signal Process., Vol. SP-40, pp. 901–914, 1992.

[25] Schmid, P., Segmentation of digitized dermatoscopic images by two-dimensional color clustering, IEEE Trans. Med. Imaging, Vol. MI-18, No. 2, pp. 164–171, 1999.

[26] Carpenter, G. A. and Grossberg, S., A massively parallel architecture for a self-organizing neural pattern recognition machine, Comput Vision, Graphics Image Process., Vol. 37, pp. 54–115, 1987.

[27] Carpenter, G. and Grossberg, S. Art-2: Self organization of stable category recognition codes for analog input patterns, Appl. Opt., Vol. 26, pp. 4919–4930, 1987.

[28] Carpenter, G. and S. Grossberg, Art-3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures, Neural Networks, Vol. 3, pp. 129–152, 1990.

[29] Grossberg, S., Embedding fields: A theory of learning with physiological implications, J. Math. Psychol., Vol. 6, pp. 209–239, 1969.

[30] Rumelhart D. E. and Hipser, D., Feature discovery by competitive learning, In: Parallel Distributed Processing, MIT Press, Cambridge, MA, pp. 151–193, 1986.

[31] Mitra, S. and Yang, S. Y., High fidelity adaptive vector quantization at very low bit rates for progressive transmission of radiographic images, J. Electron Imaging, Vol. 11, No. 4(Suppl. 2), pp. 24–30, 1998.

[32] Mitra, S., Castellanos, R., Yang, S. Y., and Pemmaraju, S., Adaptive clustering for image segmentation and vector quantization, In: Soft-Computing for Image Processing, Editors: Pal, S. K., Ghosh, A., and Kundu, M. K., eds., Springer-Verlag, New York, 1999.

[33] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., Optimization by simulated annealing, Science, Vol. 220, pp. 671–680, 1983.

[34] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., Equations of state calculations by fast computing machines, J. Chem. Phys., Vol. 21, No. 6, pp. 1087–1091, 1953.

[35] Johnston, B., Atkins, M. S., Mackiewich, B., and Anderson, M., Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI, IEEE Trans. Med. Imaging, Vol. 15, pp. 154–169, 1996.

[36] Zijdenbos, A. P., Dawant, B. M., Margolin, R. A., and Palmer, A. C., Morphometric analysis of white matter lesions in MR images: Method and validation, IEEE Trans. Med. Imaging, Vol. 13, No. 4, pp. 716–724, 1994.

[37] Hall, L. O., Bensaid, A. M., Clarke, L. P., Velthuizen, R. P., Silbiger, M. S., and Bezdek, J. C., A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain, IEEE Trans. Neural Networks, Vol. 35, pp. 672–682, 1992.

[38] Clark, M. C., Hall, L. O., Goldgof, D. B., *et al.*, MRI segmentation using fuzzy clustering-techniques, IEEE Eng. Med. Biol., Vol. 13, No. 5, pp. 730–742, 1994.

[39] Kamber, M., Collins, D. L., Shinghal, R., Francis, G. S., and Evans, A. C., Model-based 3-D segmentation of multiple sclerosis lesions in dual-echo MRI data, Proc. SPIE Visual. Biomed. Comput., Vol. 1808, pp. 590–600, 1992.

[40] Jackson, E. F., Narayana, P. A., Wolinsky, J. S., and Doyle, T. J., Accuracy and reproducibility in volumetric analysis of multiple sclerosis lesions, J. Comut. Assisted Tomog., Vol. 17, No. 2, pp. 200–205, 1993.

[41] Kapouleas, I., Automatic detection of white matter lesions in magnetic resonance brain images, Comput. Methods Programs Biomed., Vol. 32, pp. 17–35, 1990.

[42] Mitchell, J. R., Karlik, S. J., Lee, D. H., and Fenster, A., Automated detection and quantification of multiple sclerosis lesions in MR volumes of

the brain, Proc. SPIE Med. Imag. VI: Image Process., Vol. 1652, pp. 99–106, 1992.

[43] Johnston, B. G., Atkins, M. S., and Booth, K. S., Partial volume segmentation in 3-D of lesions and tissues in magnetic resonance images, Proc. SPIE Med. Imaging, Vol. 2167, pp. 28–39, 1994.

[44] Gerig, G., Kübler, O., Kikinis, R., and Jolesz, F. A., Nonlinear anisotropic filtering of MRI data, IEEE Trans. Med. Imaging, Vol. 11, pp. 221–232, 1992.

[45] Leemput, K. V., Maes, F., Vandermeulen, D., Colchester, A., and Suetens, P., Automated segmentation of multiple sclerosis lesions by model outlier detection, IEEE Trans. Med. Imaging, Vol. 20, No. 8, pp. 677–688, 2001.

[46] Kwan, R. K.-S., Evans, A. C., and Pike, G. B., MRI simulation-based evaluation of image-processing and classification methods, IEEE Trans. Med. Imaging. Vol. 18, No. 11, pp. 1085–97, 1999.

[47] Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., and Evans, A. C., BrainWeb: Online Interface to a 3D MRI Simulated Brain Database, available at: http://www.bic.mni.mcgill.ca/brainweb/.

[48] Corona, E., Mitra, S., Wilson, M., and Soliz, P., Digital stereo optic disc image analyzer for monitoring progression of glaucoma, Proc. SPIE, Vol. 4684, pp. 82–93, 2002.

[49] Corona, E., Mitra, S., Wilson, M., Krile, T., Kwon, Y. H., and Soliz, P., Digital stereo image analyzer for generating automated 3-D measures of optic disc deformation in glaucoma, IEEE Trans. Med. Imaging, Vol. 2, No. 10, pp. 1244–1253, 2002.

[50] Yang, S., King, P., Corona, E., Wilson, M., Aydin, K., Mitra, S., Soliz, P., Nutter, B., and Kwon, Y. H., Feature extraction and segmentation in medical images by statistical optimization and point operation approaches, Proc. SPIE, Vol. 5032, pp. 1676–1684 , 2003.

[51] Mitra, S., Nutter, B. S., and Krile, T. F., Automated method for fundus image registration and analysis, Appl. Optics, Vol. 27, pp. 1107–1112, 1988.

[52] Lee, D. J., Krile, T. F., and Mitra, S., Power cepstrum and spectrum techniques applied to image registration, Appl. Optics, Vol. 27, pp. 1099–1106, 1988.

[53] Sun, C., A fast stereo matching method, In: Proceedings of Digital Image Computing: Techniques and Applications, Massey University, Auckland, New Zeland, December 10–12, 1997, pp. 95–100.

[54] Ramirez, J., Mitra, S., and Morales, J., Visualization of the three dimensional topography of the optic nerve head through a passive stereo vision model, J. Electron. Imaging, Vol. 8, No. 1, pp. 92–97, 1999.

[55] Smith, P. W. and Nandhakumar, N., An improved power cepstrum based stereo correspondence method for textured scenes, IEEE Trans. Patt. Anal. Machine Intell., Vol. 18, No. 3, pp. 338–348, Mar. 1996.

[56] Pratt, W. K., Digital Image Processing, 2nd edn., Wiley-Interscience, New York, pp. 112–117, 1991.

[57] Kirbas, C. and Quek, F. K. H., Vessel extraction techniques and algorithms: A survey, In: 3rd Symposium on Bioinfomatics and BioEngineering, Bethesda, Maryland, March 2003, pp. 238–245.

[58] Kass, M., Witkin, A., and Terzoopoulos, D., Snakes: Active contour models, Int. J. Comp. Vision, Vol. 1, pp. 321–331, 1988.

[59] Osher, S. and Sethian, J. A., Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulation, JCP, Vol. 79, pp. 12–49, 1988.

[60] Chaudhuri, S. C., Katz, N., Nelson, M., and Goldbaum, M., Detection of blood vessels in retinal images using two dimensional blood vessel filters, IEEE Trans. Med. Imaging, Vol. 8, pp. 263–269, 1989.

[61] Wood, S. L., Qu, G., and Roloff, L. W., Detection and labeling of retinal vessels for longitudinal studies, In: IEEE International Conference on Image Processing, 1995, Vol. 3, pp. 164–167.

[62] Hoover, A., Kouznetsova, V., and Goldbaum, M., Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Trans. Med. Imaging, Vol. 19, pp. 203–210, 2000.

[63] Zana, F. and Klein, J. C., Robust segmentation of vessels from retinal angiography, in IEEE International Conference on Digital Signal Processing, 1997, Vol. 2, pp. 1087–1090.

[64] Tolias, Y. and Panas, S. M., A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering, IEEE Trans. Med. Imaging, Vol. 17, pp. 263–273, 1998.

[65] Zhou, L., Rzeszotarski, M. S., Singerman, L. J., and Chokreff, J. M., The detection and quantification of retinopathy using digital angiograms, IEEE Trans. Med. Imaging, Vol. 13, pp. 619–626, 1994.

[66] Rosenfeld, A., On connectivity properties of grayscale pictures, Patt. Recogn., Vol. 16, No. 1, pp. 47–50, 1983.

[67] Wang, W., Sun, C., and Chao, H., Color image segmentation and understanding through connected components, In: Proceedings of 1997 IEEE Int'l Conf. on Systems, Man, and Cybernetics, Orlando, FL, Oct. 12–15, 1997, Vol. 2, pp. 1089–1093.

[68] Castellanos, R., Castillo, H., and Mitra, S., Performance of nonlinear methods in medical image restoration, SPIE Proc. Nonlinear Image Process., Vol. 3646, 1999.

[69] Jain, A. K. and Dubes, R. C., Algorithms for Clustering Data, Prentice Hall, Englewood Cliffs, NJ, 1988.

[70] Johnson, K. A., and Becker, J. A. The Whole Brain Atlas, available at: http://www.med.harvard.edu/AANLIB/.

Chapter 7

Automatic Analysis of Color Fundus Photographs and Its Application to the Diagnosis of Diabetic Retinopathy

Thomas Walter[1] *and Jean-Claude Klein*[1]

7.1 Introduction

Medical image processing is the meeting of two sciences that behave in completely different ways. While medicine is a science where experience plays a majors role and where the practical use is evident, image processing—as a derivative of applied mathematics—is a more theoretical discipline. Hence, the conditions of this meeting need to be analyzed sophisticatedly; not everything possible to implement is useful, and not everything useful is possible to implement.

In this introductory part, we describe the biomedical context and the clinical motivation of the methods presented in this chapter.

7.1.1 The Anatomy of the Fundus

Diabetic retinopathy affects—as the name tells us—the *retina*, a layer of the back of the eyeball (also called *fundus*). In this section, we describe briefly its anatomy.

The fundus consists of three layers (see also Fig. 7.1):

[1] Centre of Mathematical Morphology, Paris School of Mines, Fontainebleau, France

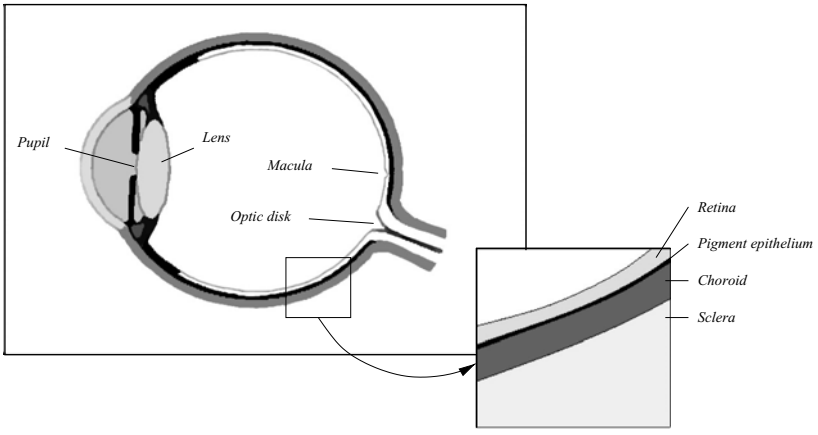


Figure 7.1: The anatomy of the eye.

- Sclera: The tough outer part of the eye.
- Choroid: A spongy layer filled with blood vessels. Its main function is to nourish the retina.
- Retina: The innermost layer of the eyeball. Place where the image created by the lens is focused and transformed into nerve impulses which are then sent to the brain via the optic nerve.

The retina itself can be divided into two layers: The photoreceptor layer and the pigment epithelium (sometimes, the latter one is introduced not as a part of the retina, but as a layer of its own). The pigment epithelium has metabolic functions; it lies between the photoreceptor layer and the choroid and it is densely packed with pigment granules.

The cells responsible for the transformation of light into nerve impulses are the rods and the cones. They are not distributed uniformly on the retina: The concentration of the cones—responsible for daylight vision—is maximal in the *macula*, the center of vision.

Once the light has been transformed into a nerve impulse, the information has to be transported to the brain. This is done by the optic nerve that enters the eye by the *optic disk* (or *papilla*). The papilla does not contain any photoreceptor; it is also called the *blind spot*.

The *retinal vessels* that nourish the retinal tissue also enter by the optic disk: Only the macula is exclusively nourished by the choroidal layer and is therefore not vascularized.

These three features—the macula, the papilla, and the vascular tree—are the main anatomical features of the retina.

The retina can be seen as an exterior part of the brain; it is highly specialized and complicated. There are many diseases that can affect this part of the eye and one of the most important is *diabetic retinopathy*.

7.1.2 Diabetic Retinopathy

Diabetic retinopathy is a severe and widespread eye disease. In fact, it is the leading cause of legal blindness for the working age population (between 20 and 64) in western countries. Diabetic retinopathy is a complication of *diabetes mellitus* and its prevalence increases with the duration of the disease. After 15 years of diabetes, the prevalence lies near 98%, so nearly all diabetic patients are affected by this disease after some time. Although not all the forms of the disease coincide with vision alteration, about 2% of the diabetic patients are blind and 10% suffer from vision loss after 15 years of diabetes [1].

This shows that diabetic retinopathy is a very frequent eye disease, but the situation shall become even worse in the future. The number of diabetics in the world is strongly increasing; a number of 300 millions of diabetic patients is expected for the year 2025. Hence, diabetic retinopathy is a major problem for an increasing number of persons, and also for the national health systems. According to [2], blindness due to diabetic eye disease produces costs of about 500 millions dollars a year in the United States.

7.1.2.1 The Evolution of Diabetic Retinopathy

Diabetic retinopathy is a silent disease, i.e. in its early stages, it is asymptomatic and vision is not altered. Vision impairment and blindness are the consequences of the complications of this disease, not of the disease itself. The starting point of the disease is the elevation of glucose in the blood which results in alterations of the vascular walls, like microaneurysms, the first unequivocal sign of the disease (see Fig. 7.2). This first abnormality causes two phenomena:

- Capillary occlusions due to modifications in the capillary walls result in local ischemia (a region in the retina which is no longer supplied with blood) accompanied by hemorrhages. If this ischemia is relatively large, it may give rise to *new vessels* which proliferate on the surface of the

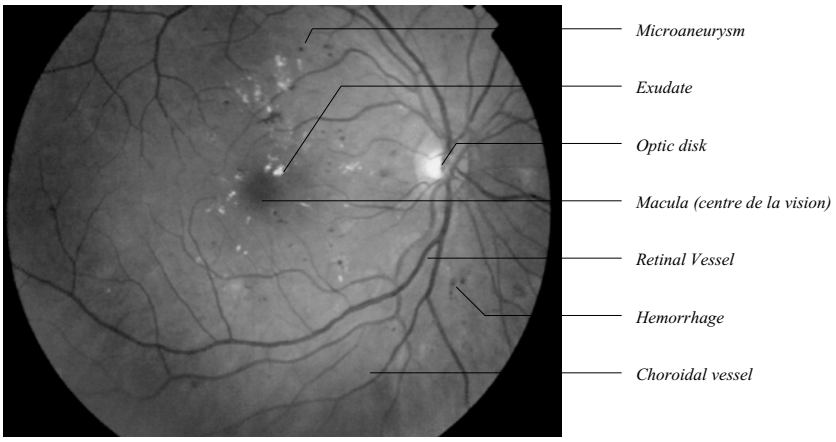


Figure 7.2: Image of the fundus.

retina. These new vessels are normally weak and may cause vitreous hemorrhages, which are one of the main reasons for irreparable vision impairment and blindness due to diabetic retinopathy.

- Because of alterations of the vascular walls, the vessels become hyperpermeable. As a consequence, extracellular liquid can pass the wall and may accumulate under the retina building a *retinal edema*. If situated in the macula, the center of vision, it is called *macular edema*; it is accompanied by exudates in the macular region (intraretinal deposits made of serum lipoproteins, see Fig. 7.2) and it represents one of the main reasons for vision loss.

Both of these two main complications can be prevented by an adapted treatment if the disease is detected early enough. Hence, early diagnosis of diabetic retinopathy is essential for the prevention of vision impairment and blindness threatening a large number of patients.

7.1.3 Application of Image Analysis to the Diagnosis of Diabetic Retinopathy

From a medical point of view, there are three main domains of algorithms which can be conceived for the improvement of diagnosis of diabetic retinopathy:

1. *Image enhancement*: Images taken at standard examinations are often noisy and poorly contrasted. Over and above that illumination is normally not uniform. Techniques improving contrast and sharpness and reducing noise are therefore required.

2. *Mass screening*: Computer-assisted mass screening is certainly the most important task to which image processing can contribute. We have already seen that the blinding complication of diabetic retinopathy can be inhibited by early treatment. However, as vision normally alters only in the later stages of the disease, many patients remain undiagnosed in the earlier stages of the disease. Hence, mass screening of all diabetic patients would help to diagnose this disease early enough. Unfortunately, this approach is not very realistic, taking into consideration the large number of diabetic patients compared to a lack to specialists. Computer assistance could make mass screening a lot more efficient.

3. *Monitoring*: Comparing images taken at different examinations allows one to evaluate a treatment or new therapeutics. However, it is a time-consuming task and open to human error. Computer-based comparison including automatic registration and evaluation of changes between images could deliver a precious tool for monitoring the disease.

We have seen in this introductory section that diabetic retinopathy is a real problem for a high number of diabetic patients and even for our health systems. We have also seen possible approaches of computer assistance that may help to overcome actual problems in its diagnosis.

Of course, giving detailed solutions to all these problems would go over the scope of this paper. After having analyzed the nature of color in fundus photographs and after having given a short introduction into mathematical morphology, a nonlinear image-processing technique our algorithms are mainly based on, we will describe in detail some algorithms within this framework: We will present an algorithm for image enhancement, algorithms for the detection of the vessels and the optic disk, and finally algorithms for the detection of characteristic lesions like microaneurysms and exudates. These segmentation algorithms are essential for computer-assisted screening and monitoring systems.

7.2 Interpretation of the Color of Retinal Images

Color images are becoming increasingly important for the diagnosis of diabetic retinopathy; their acquisition is cheap, noninvasive, and easy to perform. It is only in the last decade that they have become—due to considerable technical improvement of their acquisition—really important for the diagnosis of this disease. Before, fluorescein angiographies have been used for years. Although the latter ones still allow detection of microaneurysms—the lesion characteristic to diabetic retinopathy—with a greater sensitivity, they are invasive and costly and therefore not adapted for screening purposes.

In this section, we discuss the color content of fundus images, and we deduce a color representation which is adapted for automatic treatment.

7.2.1 The Spectral Response of the Fundus

The reflectance of the fundus strongly depends on the wavelength of the incident light. It is very high for the red light [3]; this is why fundus images seem to be "reddish." However, this does not mean that the response to red light reveals most information about the retina. In order to analyze the information content of the color, we have to analyze, more precisely, the process of reflection at and absorption in the different layers of the inner eye (see [27] for a detailed discussion).

In [4] and [3], a radiation transport model is proposed that helps understanding the color content of fundus images. The light enters the eye by the pupil and transverses the layers of the inner eye. In all layers, there is a part of this incoming light which is absorbed, a part which passes the layer, and another part which is reflected. It is this last part of the light that characterizes the perceived color. The transmission, absorption, and reflection of light with a given wavelength depend mainly on the tissue properties, mainly on the concentration of the two pigments melanin and hemoglobin.

The light in the blue spectrum is strongly absorbed by both melanin and hemoglobin. This is one of the main reasons why the reflected light does not contain a lot of blue: It is nearly entirely absorbed in the pigment epithelium layer. Furthermore, the dispersion of the light depends on its wavelength and it

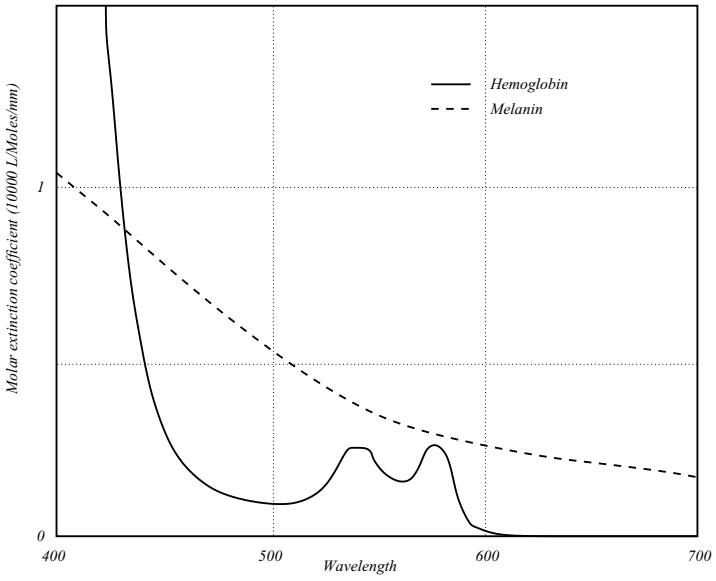


Figure 7.3: The extinction coefficient of hemoglobin and melanin depending on the wavelength of the incoming light.

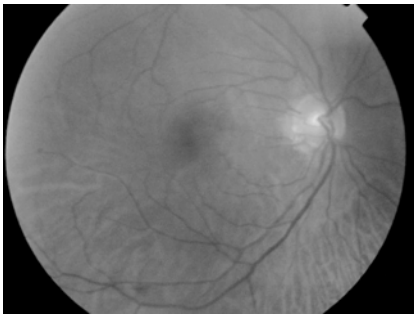
is stronger for smaller wavelength. As a consequence, opacities disturb the blue light more than light of other wavelengths.

Green light is also absorbed by the two pigments, but less than blue light. We can also observe in the Fig. 7.3 that the absorption coefficient has a peak for the green light. As a consequence, features containing hemoglobin absorb more green light than the surrounding tissue; they appear dark in images taken with green light. The green light is reflected on the pigment epithelium and it does not enter into the choroidal layer.

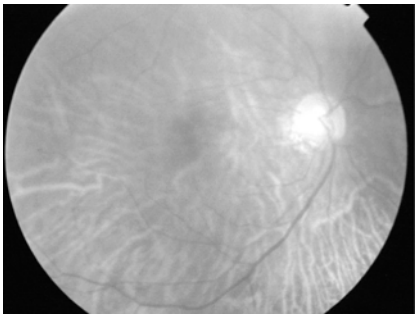
Red light, in contrary, whose absorption by hemoglobin and melanin is quite weak, penetrates deeper into the layers of the inner eye; it is mainly reflected at the sclera. Hence, the red part of the reflected light comes from the choroidal layer or the sclera; it does not contain much information about the retina itself.

7.2.2 RGB Representation

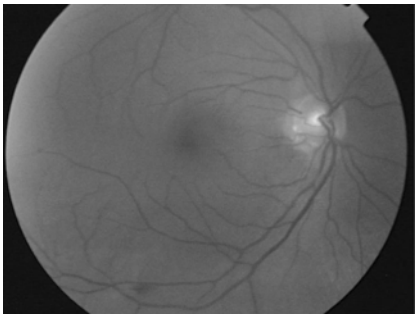
These considerations explain why the use of green light is very advantageous for the analysis of the retina, particularly for the visualization and analysis of blood containing elements. The RGB representation of the color image allows



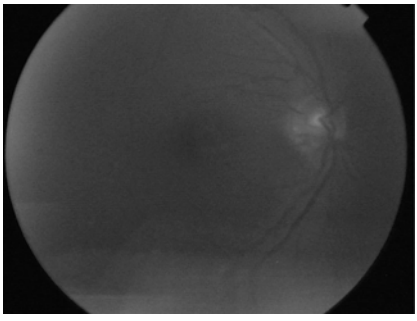
(a) A color image



(b) The red channel



(c) The green channel



(d) The blue channel

Figure 7.4: A *RGB* representation of a color image.

one to exploit this interpretation of the reflected spectrum. A color image and its decomposition into the three channels, red, green, and blue, is shown in Fig. 7.4. Of course, the given interpretation holds only approximately: The red channel is not the spectral response to red illumination, but the red part of the spectral response to illumination with white light.

The RGB representation of a color fundus image is shown in Fig. 7.4. The red channel is relatively bright and the vascular structure of the choroid is visible. The retinal vessels are also visible but less contrasted than in the green channel (compare Fig. 7.4(b) with Fig. 7.4(c)). The blue channel is noisy and contains only few information. The vessels are hardly visible and the dynamic is very poor.

This phenomenon can be observed in all retinal images we have studied (about 200) with one difference: Sometimes, the blue channel contains information, sometimes, it does not. Indeed, the quality of the blue channel of the RGB

color space depends on the age of the patient and on the yellowing of the cornea; the cornea can be understood as a filter that screens out ultraviolet radiation. With the age, the cut-off frequency moves toward the blue, and even blue light can no longer pass.

This interpretation of the color content of fundus images privileges the use of the RGB color space, for the channels have a physical meaning. We have compared qualitatively the green channel of 30 fundus images with channels of other color spaces (HSV, HLS, Lab, Luv, principal components) and for all images, the green channel was better contrasted than any other channel (at least concerning all blood-containing features). Another advantage of the use of the green channel is that the choroidal vessels do not appear at all, whereas they do appear in the luminance channel for instance, for it is a combination of the three channels R, G, and B. This is why, we work mainly on the green channel.

7.3 Basic Morphological Operators

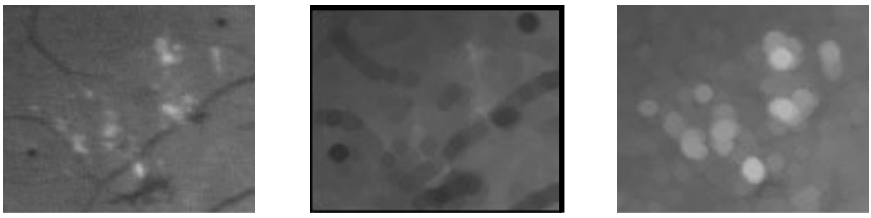
Mathematical morphology is a nonlinear image-processing technique. The interested reader may see [5] for an exhaustive discussion and [6] for a comprehensive introduction.

The basic operators presented in this chapter deal with two-dimensional discrete images (defined on $E \subset \mathbb{Z}^2$). Binary images are defined as subsets of E and gray scale images as functions $f : E \to T$, with $T = \{t_{\min}, \ldots, t_{\max}\}$ the set of gray levels.

7.3.1 Erosion and Dilation

Many operators in mathematical morphology are based on the use of a small "test-set" B called structuring element (SE). Its shape and size can be chosen in accordance with the segmentation or filtering task.

In order to calculate the morphological erosion of a binary image A, we test for each point x if the structuring element centered in x fits completely into A. If this is the case, x belongs to the eroded set εA. The dilation can be seen as an erosion of the background.



(a) Original (b) Erosion (c) Dilation

Figure 7.5: Erosion and dilation with a circular SE of a retinal image (detail).

The gray level dilation/erosion substitutes the value $f(x)$ by the maximum/minimum of f for all the pixels contained in the translated SE B_x:

$$\begin{aligned}\left[\varepsilon^B(f)\right](x) &= \min_{b\in B} f(x+b)\\ \left[\delta^B(f)\right](x) &= \max_{b\in B} f(x+b)\end{aligned} \tag{7.1}$$

In Fig. 7.5, the effect of these operations are shown. We see that the erosion enlarges dark details and reduces bright ones and the dilation enlarges bright details and reduces dark ones.

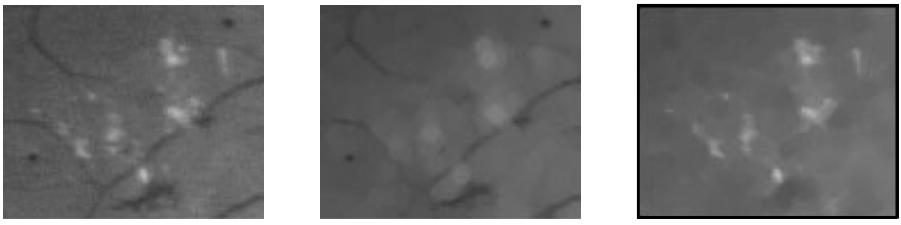
7.3.2 Opening and Closing

Morphological openings γ^B and closings ϕ^B are the consecutive application of erosion and dilation:

$$\begin{aligned}\gamma^B(\cdot) &= \delta^{\check{B}}\varepsilon^B(\cdot)\\ \phi^B(\cdot) &= \varepsilon^{\check{B}}\delta^B(\cdot)\end{aligned} \tag{7.2}$$

with $\check{B} = \{-b\}$ the transposed structuring element.

In order to understand the behavior of gray scale openings and closings, it is useful to consider the image f as a topographic surface: Pixels with low gray-levels correspond to valleys, pixels with high gray-levels correspond to mountains. A morphological opening removes all bright features that cannot contain the structuring element; "it razes the elevations." A morphological closing removes all dark features that cannot contain the structuring element; "it fills the depressions in the surface." The opening and closing of a retinal image are shown in Fig. 7.6: The opening removes the bright patterns (exudates) without enlarging the dark ones (Fig. 7.6(b)). The closing removes the dark patterns



(a) Original (b) Opening (c) Closing

Figure 7.6: Opening and closing of a retinal image (detail) with a circular SE.

(hemorrhages, vessels, microaneurysms) without enlarging the bright ones; but the shape of the bright patterns are altered, the spaces between the exudates are also filled (Fig. 7.6(c)).

Furthermore, any increasing, idempotent, and antiextensive transformation is called *algebraic opening*. Any increasing, idempotent, and extensive transformation is called *algebraic closing*.

7.3.3 The Morphological Reconstruction

In order to avoid these alterations of contours, like the ones caused by openings and closings shown in Fig. 7.6, the morphological reconstruction can be used.

The morphological reconstruction by dilation works with two images: a marker f and a mask g. The marker image is dilated, then the point-wise minimum with g is calculated, then the result is dilated once again, and so on. This process is iterated until idempotence. Let $\delta_g^{(1)} f = \delta^B f \wedge g$. The morphological reconstruction can then be written as:

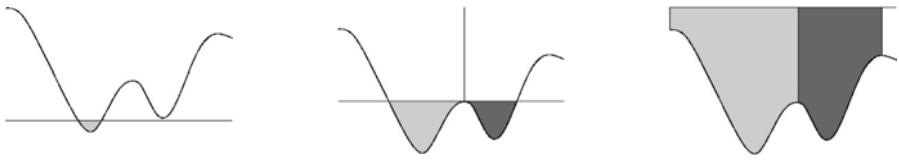
$$R_g(f) = \delta_g^{(\infty)}(f) \quad \text{with } \delta_g^{(n)}(f) = \delta_g^{(1)}\left[\delta_g^{(n-1)}(f)\right] \tag{7.3}$$

With $\varepsilon_g^{(1)} f = \varepsilon^B f \vee g$, the reconstruction by erosion can be defined analogously:

$$R_g^*(f) = \varepsilon_g^{(\infty)}(f) \quad \text{with } \varepsilon_g^{(n)}(f) = \varepsilon_g^{(1)}\left[\varepsilon_g^{(n-1)}(f)\right] \tag{7.4}$$

7.3.4 The Watershed Transformation

One of the most powerful tools for image segmentation in mathematical morphology is the watershed transformation [7]. A gray-level image f is interpreted as a topographic surface and a flooding process is simulated starting from the



(a) Starting with the local minima

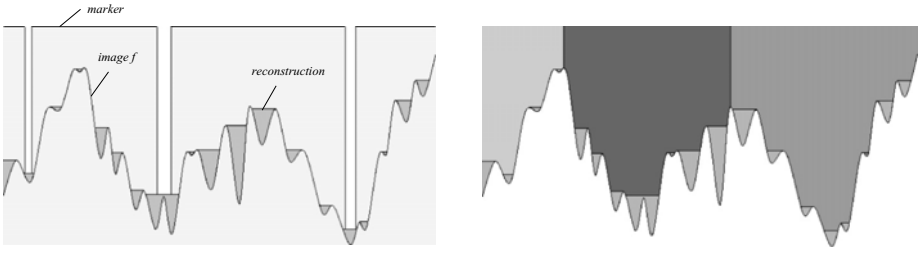
(b) Two lakes meeting

(c) Watershed transformation

Figure 7.7: The watershed transformation.

local minima ("sources"). The flooding level s is the same for the whole image; all pixels with a gray level value lower than s belong therefore to a "lake" (see Fig. 7.7(a)). When two lakes meet, a "wall" is built between the two lakes, i.e., the pixel where the two lakes meet forms part of the watershed line $WS(f)$ (see Fig. 7.7(b)). The whole image is flooded in this way giving an image that contains the watershed line $WS(f)$ and as many regions as local minima in the original image f (see Fig. 7.7(c)). These regions are called catchment basins CB_i in analogy to their topographic interpretation.

The presence of many minima dues to the noise present in real images results in over-segmentation. The number of minima can be reduced before calculating the watershed transformation by means of the morphological reconstruction. Therefore, we calculate a marker image m, which takes the value $f(x)$ for all the "marked pixels" and $t_{\max}$ elsewhere (see Fig. 7.8(a)). Then, we calculate the reconstruction by erosion $R^*_f(m)$, i.e., we remove ("fill") all not marked minima (Fig. 7.8(a)). For this modified image, the watershed transformation gives a more persistent result (Fig. 7.8(b)).



(a) An image f, a marker m (in gray), and the reconstruction by erosion

(b) The watershed transformation

Figure 7.8: The watershed transformation controlled by a marker m.

7.4 Contrast Enhancement and Correction of Nonuniform Illumination

There are three systematic problems that occur in nearly all segmentation tasks of color fundus images:

- Nonuniform illumination
- Poor contrast
- Noise

For the attenuation of noise, we cannot propose filters that are applicable in general, because the exigency on such filters depends on the segmentation task. If big features are to be detected (e.g. the macula), strong filters can be used whereas algorithms dedicated to the detection of small details (e.g. microaneurysms) must rely on filters that preserve even small dark details.

In this section, we present an algorithm for contrast enhancement and shade correction. First, we propose a simple global contrast enhancement operator. Applying this operator locally enhances the contrast and corrects nonuniform illumination in one step.

7.4.1 Polynomial Contrast Enhancement

Let $f : E \to T$ be a gray-level image with $T = \{t_{\min}, \ldots, t_{\max}\} \subset \mathbb{R}$ a set of rational numbers. Let $U = \{u_{\min}, \ldots, u_{\max}\} \subset \mathbb{R}$ be a second set of rational numbers. An application

$$\begin{aligned} \Gamma : T &\to U \\ u &= \Gamma(t) \end{aligned}$$

is called *gray-level transformation*.

For convenience, the gray-level transformation is constructed in such a way that it assigns $\frac{1}{2}(u_{\min} + u_{\max})$ to the mean value μ_t of the original image f. Instead of t and u, we consider in the following the variables τ and ν defined by:

$$\begin{aligned} \tau &= t - \mu_t \\ \nu &= u - \frac{1}{2}(u_{\min} + u_{\max}) \end{aligned} \tag{7.5}$$

A polynomial gray-level transformation can then be defined as follows:

$$\nu = \Gamma^*(\tau) = \begin{cases} a_1 \cdot (\tau - \tau_{\min})^r + b_1 & \text{if} \quad \tau \leq 0 \\ a_2 \cdot (\tau - \tau_{\max})^r + b_2 & \text{if} \quad \tau \geq 0 \end{cases} \tag{7.6}$$

with the parameters r, a_1, a_2, b_1, and b_2. The parameter r can be chosen freely, the other parameters are determined in order to assure that the transformation Γ is continuous and that the resulting image covers the whole gray-level range (from $u_{\min}$ to $u_{\max}$). These conditions can be expressed by

$$\begin{aligned} \Gamma^*(\tau_{\min}) &= \nu_{\min} \\ \lim_{\tau \to 0^-} \Gamma^*(\tau) &= 0 \\ \lim_{\tau \to 0^+} \Gamma^*(\tau) &= 0 \\ \Gamma^*(\tau_{\max}) &= \nu_{\max} \end{aligned} \tag{7.7}$$

and with Eq. (7.6) we obtain for the parameters a_1, a_2, b_1, and b_2:

$$\begin{aligned} a_1 &= \frac{-\nu_{\min}}{(-\tau_{\min})^r} = \frac{\frac{1}{2}(u_{\max} - u_{\min})}{(\mu_t - t_{\min})^r} \\ a_2 &= \frac{-\nu_{\max}}{(-\tau_{\max})^r} = \frac{-\frac{1}{2}(u_{\max} - u_{\min})}{(\mu_t - t_{\max})^r} \\ b_1 &= \nu_{\min} = \frac{1}{2}(u_{\min} - u_{\max}) \\ b_2 &= \nu_{\max} = \frac{1}{2}(u_{\max} - u_{\min}) \end{aligned} \tag{7.8}$$

and finally, for $u = \Gamma(t)$:

$$u = \Gamma(t) = \begin{cases} \frac{\frac{1}{2}(u_{\max} - u_{\min})}{(\mu_t - t_{\min})^r} \cdot (t - t_{\min})^r + u_{\min} & \text{if} \quad t \leq \mu_t \\ \frac{-\frac{1}{2}(u_{\max} - u_{\min})}{(\mu_t - t_{\max})^r} \cdot (t - t_{\max})^r + u_{\max} & \text{if} \quad t > \mu_t \end{cases} \tag{7.9}$$

The corresponding graph is shown in the Fig. 7.9 for different μ_t. The resulting transformation is not symmetric to the point $(\mu_t, \frac{1}{2}(u_{\max} + u_{\min}))$.

With r, we can control the strength of the contrast enhancement. For $\mu_t = \frac{1}{2}(t_{\min} + t_{\max})$, we obtain a linear contrast stretching operator for $r = 1$. For $r \to \infty$, we obtain a threshold operation with the thresh μ_t.

If this operator is applied to the whole image as a global contrast operator, the result is not satisfying due to the nonuniform illumination. In fact, the proposed gray-level transformation does not enhance the contrast for any subset of T, but only for subsets for which $\frac{\partial u}{\partial t} > 1$. For instance, the contrast of a dark detail situated in a dark region may even be attenuated.

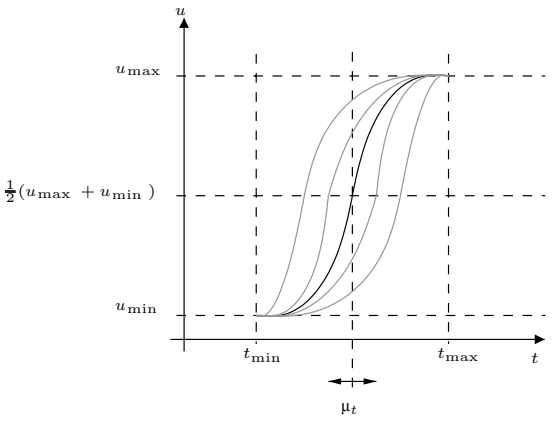


Figure 7.9: The graph of the gray level transformation for different μ_t.

7.4.2 Contrast Enhancement and Shade Correction

In order to enhance the contrast all over the image independently from local illumination changes, we propose a shade correction operator based on the gray-level transformation shown in the preceding section.

Shade correction: A shade correction operator tries to remove the background information from an image. This is done by calculating a background approximation (for example with a low pass filter) and by subtracting it from the image. In order to avoid negative values, a constant is usually added:

$$[\mathrm{SC}(f)](x) = f(x) - [A(f)](x) + c \tag{7.10}$$

In the corrected image, the gray-level values depend only on the difference between the original value and the background approximation.

The local contrast enhancement operator: In order to obtain a shade correction operator, which also enhances the contrast, we apply the gray-level transformation from Eq. (7.9) locally, i.e. we substitute the global mean μ_t by a local background approximation.

One possibility is to calculate the mean value of f within a window W centered in the pixel x:

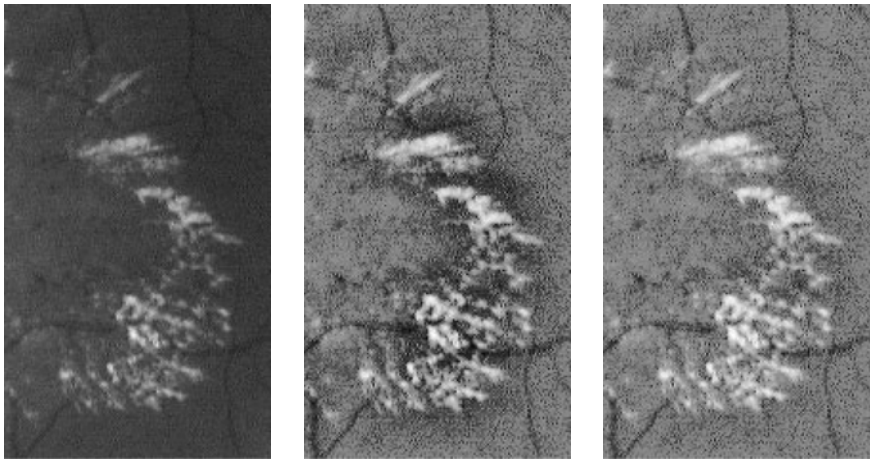
$$\mu_t^W(x) = \frac{1}{N_W} \sum_{\xi \in W(x)} f(\xi) \tag{7.11}$$

In this way, a contrast operator is obtained for which the transformation parameters depend on the mean value of the image in a window of a certain size. Hence, it is a shade correction and contrast-enhancement operator.

However, for pixels close to bright features, the background approximation may be biased by blurred bright objects. Indeed, we observe a "darkening" close to bright objects as the papilla or exudates (see Fig. 7.10b). This darkening is a real problem for segmentation algorithms, because these regions may then be confused with vessels, hemorrhages, or microaneurysms. Therefore, we propose to calculate the local mean value on a filtered image, where all these bright features have been removed. We have seen that the morphological opening removes bright features from an image (see section 7.3). However, we found it advantageous to apply an area opening γ_λ^a [8] rather than a morphological one. Instead of using a SE, γ_λ^a removes all bright objects if their area (number of pixels) is smaller than λ. The shade-correction operator can then be written as

$$[SC(f)](x) = \begin{cases} \frac{\frac{1}{2}(u_{\max}-u_{\min})}{(\mu^W_{\gamma_\lambda^a}(x)-t_{\min})^r} \cdot (t-t_{\min})^r + u_{\min} & \text{if} \quad t \leq \mu^W_{\gamma_\lambda^a}(x) \\ \frac{-\frac{1}{2}(u_{\max}-u_{\min})}{(\mu^W_{\gamma_\lambda^a}(x)-t_{\max})^r} \cdot (t-t_{\max})^r + u_{\max} & \text{if} \quad t \geq \mu^W_{\gamma_\lambda^a}(x) \end{cases} \tag{7.12}$$

The results obtained by the application of this operator are shown in Fig. 7.10 and Fig. 7.11.

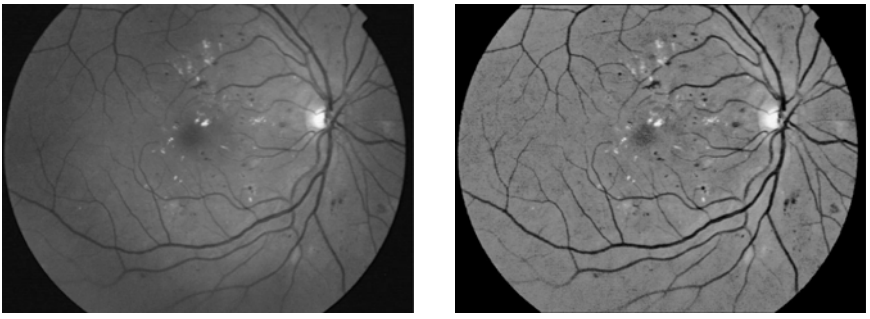


(a) Detail of the green channel of a fundus image containing hard exudates

(b) The shade correction operator with the local mean value as background approximation

(c) The final shade correction and contrast enhancement operator

Figure 7.10: The effect of filtering of the background approximation.



(a) Original image (b) The result for $r = 3$

Figure 7.11: A result for shade correction and contrast enhancement.

The shade correction of retinal images is a prerequisite of several algorithms, for example, the detection of microaneurysms shown in section 7.6.1 and the detection of vessels shown in the next section.

7.5 The Detection of Anatomical Structures in Color Fundus Images

The main anatomical features in fundus photographs are—as it has been explained in section 7.1—the vascular tree, the optic disk, and the macula. In the following subsections we present methods for the detection of the vessels and the papilla. An algorithm for the localization of the macula can be found in [9].

7.5.1 The Detection of the Vascular Tree by Means of the Watershed Transformation

In this section, we present a method for the detection of the vascular tree in color images of the human retina. This algorithm is quite general; only few information specific for retinal images is used. It can therefore be used for the extraction of elongated features in other types of images.

7.5.1.1 Motivation

Detecting the vascular tree is essential for the analysis of fundus images. The structure of the vascular tree gives useful information for other feature or lesion

detection algorithms (e.g., optic disk, macula, hemorrhages). Over and above that, it delivers landmarks for image registration.

7.5.1.2 Properties

Vessels are elongated features, much longer than, thick, reddish, and darker than the background. They enter the retina by the optic disk and go all over the retina forming the vascular tree.[2] With increasing distance from the optic disk, the vessels become thinner and their contrast decreases. Contrast and color of vessels vary considerably from one image to another. Even in the same image, there may be color differences, as color depends on the vessel type (artery or vein), its diameter (the amount of hemoglobin that is transported), and the illumination of the retinal region where the vessel is situated.

The width of the thickest vessels is almost constant for all images taken with the same angle and the same resolution; we can state that all vascular structures in fundus images are thinner than a parameter λ (which depends on resolution and angle of the image).

As we have seen in section 7.2, vessels appear best contrasted in the green channel f_g of the color images; our algorithm for vessel detection is exclusively based on the use of this channel. The main difficulties we have to deal with are as follows:

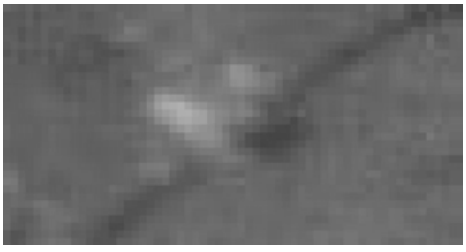
- Often, retinal images are low contrasted and corrupted by noise. As a consequence, vessel contours are not well defined, and not all vessel pixels have a lower gray level than all the background pixels. However, the mean gray level on the vessel is lower than the mean gray level on the background (see also Fig. 7.12(a)).

- The vascular tree may be interrupted by the presence of lesions (as shown in Fig. 7.12(b)) or noise.

- The presence of exudates is a source of false detections, as the spaces between exudates have sometimes properties similar to vessels in terms of luminosity, width, and connectivity (see Fig. 7.13(a)).

[2] The vascular tree as it appears in color images, is not a "tree" in the topological sense, as veins and arteries usually cross each other. It is more like a "net" of piecewise linear structures.

(a) Two vessels corrupted by noise



(b) A part of a vessel deconnected from the rest of the vascular tree by an exudate

Figure 7.12: Main problems in vessel detection.

- The presence of hemorrhages adds another source of false positives, as they have the same color. If they are connected to the vascular tree, they may be hard to distinguish from the vessels (see Fig. 7.13(b)).

7.5.1.3 State of the Art

There is a large variety of algorithms for vessel detection in retinal images. In most of these algorithms, vessels are modeled like piecewise linear segments with a Gaussian profile. Using linear or morphological filtering, features with this property are enhanced, other features are attenuated. This strategy has been proposed by Chaudhuri in [10] (linear filters) and by Zana in [11] (morphological filters). Drawbacks of these methods are the computational complexity due to directional filtering and some systematic errors on the borders of bright features (like the optic disk or hard exudates).

(a) Hard exudates close to the vascular tree

(b) Hemorrhages close to vessels

Figure 7.13: Main reasons for false detections.

Tracking algorithms are the second important group of vessel detection methods. These algorithms use the connectivity of the vascular tree as a main property. This is, in many images, not acceptable, particularly if lesions are present (see the Fig. 7.13(a)). Hence, this kind of approach must rely on good markers; then, tracking algorithms can, in our opinion, be powerful in detecting the vessel borders, but they are not adapted to detect the vascular structure.

7.5.1.4 The Algorithm

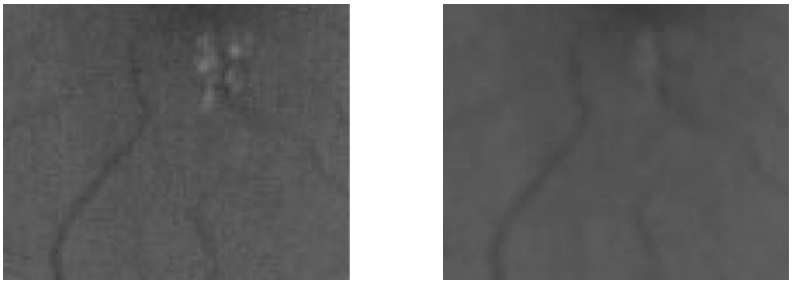
In this section, we present a new method for the detection of vessels in fundus images. The main idea is to detect thin structures in gray-scale images by evaluating the local contrast along watershed lines. This algorithm can also be applied to other problems where thin structures are to be found.

Prefiltering: As we can see in the Fig. 7.13(a), spaces between hard exudates are a systematic source of false positives for vessel detection algorithms. In order to remove small exudates, the prefiltered image p is calculated as follows:

$$p = \gamma_\lambda^a f_\mathrm{g} \tag{7.13}$$

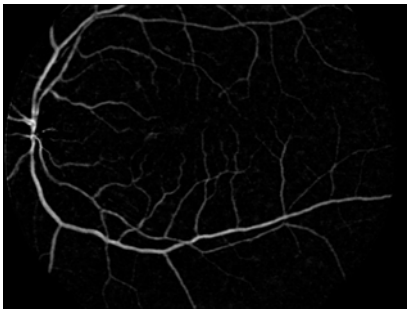
with f_g the green channel and γ_λ^a the area opening with the parameter λ. The result of this prefiltering step is shown in Fig. 7.14(b). One may notice that this filter is not very restrictive, the borders of the different features present in the image are not altered, but the small exudates are removed.

Extraction of dark details: Vessels appear as dark features in the green channel of a color image, their maximal width is known and does not vary with the image (as far as the resolution is the same). As we have seen in section 7.3,

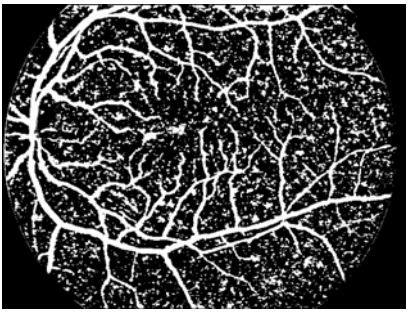


(a) The green channel (b) The prefiltered image

Figure 7.14: The prefiltering step: Small exudates are removed.



(a) The top-hat transformation of the prefiltered image



(b) An approximation of the vascular tree

Figure 7.15: Top-hat transformation and approximation of the vascular tree.

vessels can be removed from this image by means of the morphological closing with an appropriate size s_1 (see also Fig. 7.6(c)). Calculating the difference to the original gives all the dark details that cannot contain the SE:

$$\vartheta p = \phi^{s_1 B}(p) - p \tag{7.14}$$

In the top-hat image ϑp (shown in Fig. 7.15(a)), vessels appear as bright features, elongated and connected. However, because of contrast differences between retinal images and between different vessels in one image, only a raw approximation of the vascular tree can be found by means of simple threshold techniques, as shown in Fig. 7.15(b). In our example, the vessels are obtained by an area threshold T_K, proposed in [12]: The threshold is chosen in such a way that the resulting binary image contains at least K pixels.

Extraction of the crest lines: Considering the image shown in Fig. 7.15(a) as a topographic surface, we can notice that the vessels correspond to the crest lines in this image. An excellent tool for finding the crest lines in a gray-scale image has been presented in section 7.3: the watershed transformation. The strategy is to first find a good marker, then calculate the watershed transformation, and in the final step apply a contrast criterion in order to distinguish real vessels from false detection.

The usual technique to obtain a good segmentation result using the watershed transformation is to use markers (see section 7.3), i.e., the image is flooded only from "important" minima, the others are filled by means of the morphological reconstruction. Here, the markers must be chosen in such a way that the watershed line coincides with the vessels. It is, therefore, very important that

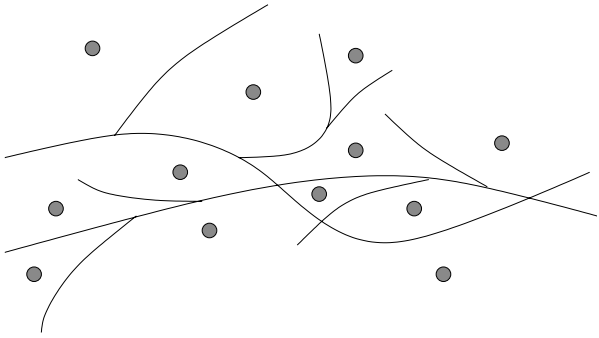


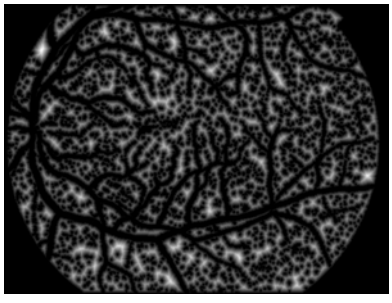
Figure 7.16: An ideal marker image (gray circles).

we mark all the "valleys" that are completely or partially surrounded by the crest lines. Such a marker is shown in the Fig. 7.16.

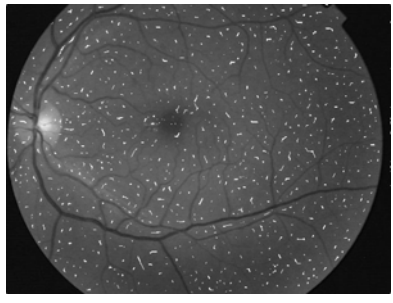
In order to obtain such a marker, we determine the points having maximal distance from the approximation shown in Fig. 7.15(b). In a first step, we fill the small "holes" of the thresholded image by a surface closing of small size, i.e., we remove all "holes" having less than 5 pixels, and then we invert the result and we determine the local maxima of the distance function:

$$\begin{aligned} m_1 &= \left[\phi_\lambda^a T_K(p)\right]^c \\ m(x) &= \begin{cases} f(x) & \text{if } x \in \mathcal{M}ax\{D(m_1)\} \\ t_{\max} & \text{if not} \end{cases} \end{aligned} \tag{7.15}$$

The distance function is shown in Fig. 7.17(a), its maxima superposed to the original image in Fig. 7.17(b). Of course, the presence of dark noise and features

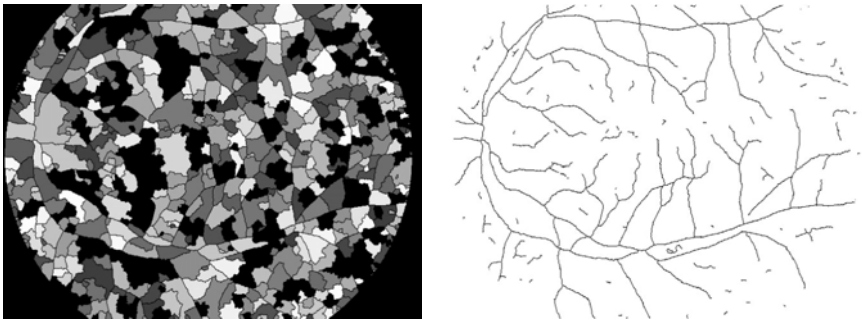


(a) The distance image of the inverted approximation



(b) The marker image (here superposed to the green channel of the original image)

Figure 7.17: A marker for vessel detection.



(a) The watershed line and the catchment basins

(b) The application of the contrast criterion

Figure 7.18: The watershed line and the result of the application of the contrast criterion.

in the original image may produce a lot of spurious objects in the approximation image. As a consequence, there are more markers than necessary, but the number of minima has been significantly reduced, and the watershed line can now be determined:

$$\mathrm{WS}_m(f) = \mathrm{WS}(R^*_f(m)) \tag{7.16}$$

Evaluation of the local contrast: The result of the watershed transformation is shown in Fig. 7.18(a). We note that on the one hand nearly all vessels coincide with a branch of the watershed line (WSL), but on the other hand, not all the branches correspond to vessels. Indeed, the high number of false positives (i.e., parts of the WSL that do not correspond to vessels) is a consequence of the fact that the WSL delimits the catchment basins. Hence, if a region is not completely enclosed by vessels, there is necessarily a branch of the WSL that does not correspond to a vessels. The nonideal marker adds even more false positives, for we do not have exactly one marker per entirely or partially enclosed region.

In order to remove these false positives, we have to analyze the WSL. We distinguish

- *bifurcation points* (BIF): all points of the WSL that have more than two neighbors.
- *branches*: all connected components of WSL\BIF(WSL). We call $F_{i,j}$ the branch being the frontier between the two catchment basins CB_i and CB_j.

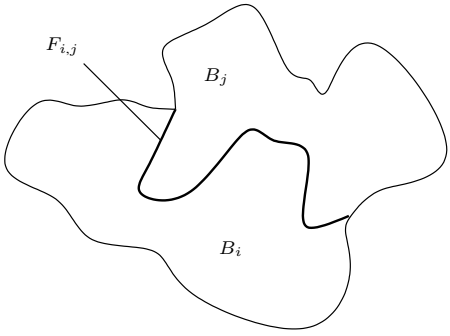


Figure 7.19: Two catchment basins BV_i and BV_j and the frontier $F_{i,j}$ between them.

In the top-hat image, vessels appear brighter than the background (brighter than the adjacent regions) and changes in gray-level on the vessels are slow. Let us now consider two catchment basins CB_i and CB_j and the frontier $F_{i,j}$ between them (see also Fig. 7.19). If $F_{i,j}$ corresponds to a vessel, the mean gray-level value of the top-hat image on $F_{i,j}$ must by higher than the mean gray level on the two catchment basins. Let ϑp be the top-hat image and $\#A$ the number of pixels of the set A. We can then write the first criterion c_1:

$$\begin{aligned}
\mu_{F_{i,j}} &= \frac{1}{\#\{F_{i,j}\}} \sum_{x \in F_{i,j}} p_2(x) \\
\mu_{BV_i} &= \frac{1}{\#\{BV_i\}} \sum_{x \in BV_i} p_2(x) \\
c_1(F_{i,j}) &= \frac{1}{2} \left((\mu_{F_{i,j}} - \mu_{BV_i}) + (\mu_{F_{i,j}} - \mu_{BV_j}) \right)
\end{aligned} \tag{7.17}$$

Evaluating the contrast criterion c_1, all the false branches not coinciding with a dark detail extracted by the top-hat are removed. However, the result is not yet satisfying, because there are still false positives that are due to some small, not connected dark details like hemorrhages close to vessels producing also a quite high value for c_1. In order to remove these false positives from the segmentation result, we have to take into consideration the local gray-level variation on the branch:

$$\begin{aligned}
\sigma_{F_{i,j}} &= \frac{1}{\#\{F_{i,j}\} - 1} \sum_{x \in F_{i,j}} (p_2(x) - \mu_{F_{i,j}})^2 \\
c_2(F_{i,j}) &= c_1(F_{i,j}) - \alpha \cdot \sigma_{F_{i,j}}
\end{aligned} \tag{7.18}$$

with α a weighting coefficient. With this enhanced contrast criterion, it is quite

simple to distinguish between vessels and false positives:

$$V_1 = \bigcup_{i,j} F_{i,j} \quad \text{with} \quad c_2(F_{i,j}) > \beta \tag{7.19}$$

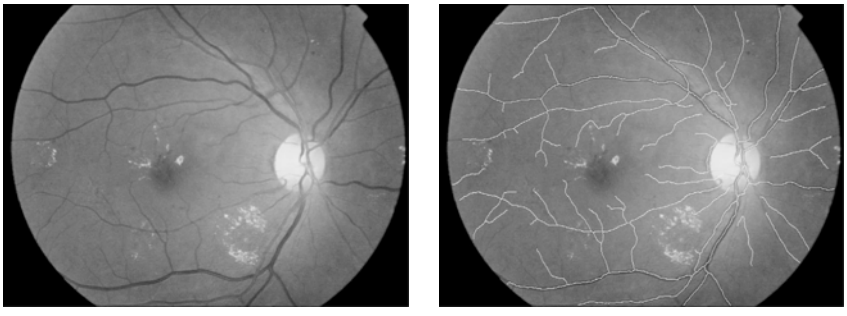
The result V_1 is shown in Fig. 7.18(b). We see that there are still small false positives. In fact, they are so small that the criterion c_2 has no meaning. Therefore, we remove all the connected components of V_1 that contain less than λ pixels (we chose $\lambda = 30$):

$$V = \gamma_\lambda^a(V_1) \tag{7.20}$$

With this technique, we obtain very satisfying results, if the images do not contain larger exudates that have not been removed by the prefiltering step. Indeed, the spaces between exudates form small channels that are quite similar to vessels. One possibility is to calculate the mean gray level for the branches in the shade corrected image $\mathrm{SC}_{\mathrm{norm}}$ and to use it as a complementary information: Only if the mean gray level is lower than a certain threshold, the branch is accepted. In this way, many false positives dues to exudates can be removed.

7.5.1.5 Results

The algorithm has been tested on sixty 640×480 fundus photographs taken with a Sony color video 3CCD camera on a Topcon TRC 50 IA retinograph. These images have not been used for the development of the algorithm. We asked an ophthalmologist to mark false detections and missed vessels on the result (*a posteriori* evaluation). We obtained a sensitivity of 83% and a predictive value of 97% an example is shown in Fig. 7.20.



(a) Original image (containing exudates)

(b) Segmentation result

Figure 7.20: A result of the vessel detection algorithm.

This kind of evaluation is certainly not the best method, as the expert is influenced by the result of the algorithm. However, vessels are clearly visible and an expert will always be able to mark them; the same holds for false positives. Over and about that, if an expert marks all vessels, it is far from being sure that he will not miss some of them, because this is a boring and time-consuming task.

7.5.2 The Detection of the Optic Disk

7.5.2.1 Motivation

The optic disk (or papilla) is one of the main features of the retina, its detection is essential for a system of automatic analysis of retinal images; it is the prerequisite for other segmentation algorithms (exudates, macula).

In the context of diagnosis of the glaucoma, the detection of and measures on the optic disk may also be of great importance. Hence, an algorithm of automatic detection of the optic disk is required.

7.5.2.2 Properties

The optic disk is the entrance of the optic nerve and the vessels into the retina. It is situated on the nasal side of the macula and it does not contain any photoreceptor: It is also called the *blind spot*. In color fundus photographs, the optic disk appears as a big bright spot of circular or elliptical shape, interrupted by the outgoing vessels. Its size varies from patient to patient, but its diameter is always comprised between 40 and 60 pixels in 640×480 images. The optic disk is characterized by a strong contrast between outgoing vessels and the bright color of the optic disk itself.

Unfortunately, this description is not valuable for all images: Sometimes, the contours are not clearly visible, the color tends more to a pale white, and there may be other regions in the image which are as bright or even brighter than the optic disk (due to nonuniform illumination or the presence of exudates).

7.5.2.3 State of the Art

In [13], the authors localize the optic disk using the high contrast between the papilla and the outgoing vessels. This method fails if there are exudates in the image.

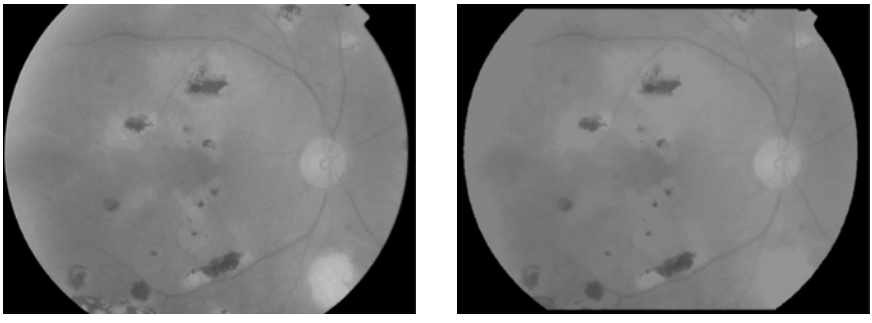
In [14], the authors use an area threshold for localization of the papilla, the Hough transform for the detection of its contours. The Hough transform is also used by [15]. The main problems that have been stated are low contrast and the case where its shape does not correspond to a circle (for example, if the optic disk is situated on the border of the image).

In [16], the authors use a template matching approach for the localization of the optic disk. The problem with this approach is the size variability of the papilla between different images and the presence of large accumulations of exudates.

7.5.2.4 The Algorithm

The presented algorithm can be subdivided into two parts: the localization and the detection of the contours of the optic disk. First versions of this algorithm have been presented in [17, 18].

Localization: As the optic disk belongs to the brightest parts of the image, the idea to apply an area threshold in order to find at least a part of the optic disk may work well, if there do not exist large accumulations of exudates or other bright regions. The atrophy in Fig. 7.21(a), for example, corresponds to a yellow spot and its size and shape are comparable to the one of the optic disk. Before we can apply a threshold to the image, it is therefore necessary to remove these bright features. This can be done using the vascular tree we have already detected: As the optic disk is the entrance of the vessels into the retina,



(a) The luminance channel of a retinal image containing an atrophy

(b) The morphological reconstruction using the vascular tree

Figure 7.21: The elimination of bright features.

it must be connected to a dilated version of the main branches of the vascular tree:

$$v(x) = \begin{cases} t_{\max} & \text{if} \quad x \in V \\ t_{\min} & \text{if} \quad x \notin V \end{cases}$$
$$l_1 = R_{f_l}\left(\delta^{sB} v \wedge f_l\right) \quad \text{with } s = 5 \tag{7.21}$$

It is recommended not to use the complete vascular tree V, but only the main branches that can be extracted easily by applying a stronger contrast criterion in the algorithm presented in section 7.5.1.

The effect of this filtering is shown in Fig. 7.21: The atrophy present in the image (a) is removed in (b), the optic disk stays nearly entirely unchanged by the reconstruction. Using the methods presented in [14, 17, 18], the localization algorithm would have failed in this case.

Now, we can assume that the optic disk belongs to the brightest elements of the image, and the application of an area threshold should give a part of the optic disk:

$$L_1 = T_{[\alpha, t_{\max}]}(l_1) \quad \text{with } \alpha \text{ such that } \#L_1 \geq K \tag{7.22}$$

L_1 normally contains more than one connected component: A part of the optic disk, some noise, and eventually other bright features connected to the vascular tree. The latter ones are normally exudates of small size. Hence, it is sufficient to choose the connected component with the largest surface to obtain a part of the optic disk:

$$L \in \mathcal{C}(L_1) \quad \text{with } \forall A \in \mathcal{C}(L_1) : \#L \geq \#A \tag{7.23}$$

The center of the (only) connected component of L can be seen as the approximative center c of the optic disk and is used for the detection of the contours described in the following paragraph.

Detection of the contours: The contours of the optic disk appear under the best contrast in the red channel f_r of the color image. Unfortunately, the red channel is sometimes saturated and cannot be used. In this case, we propose to work on the luminance channel f_l. The first step is to determine if the red channel is saturated or not. Let c be the approximative center determined in the localization step of the algorithm, f_r a subimage of the red channel centered in c, and $t_{\max}(f_r)$ the maximal gray-level value within this subimage. We define the

(a) The luminance channel

(b) The biggest particle of the threshed image

(c) The distance function of the particle

(d) The gradient image with the superposed marker

(e) The result of the watershed algorithm

(f) The segmentation result

Figure 7.22: The steps of the algorithm for the detection of the contours.

gray-level saturation S_α[3]:

$$S_\alpha = \frac{\#T_{[t_{\max}(f_r)-\alpha, t_{\max}(f_r)]}(f_r)}{\#T_{[0, t_{\max}(f_r)]}(f_r)} \tag{7.24}$$

This measure determines the percentage of pixels in the subimage whose gray level is larger than $t_{\max}(f_r) - \alpha$. If this percentage is too high, the channel is saturated and does not contain any exploitable information. We use the red channel, if for $\alpha = 30$, $S_\alpha < 0.5$ (this has been found experimentally), if not, the luminance channel is used. We call the used channel f_c in the following.

For finding the contours of the optic disk, we shall make use of the watershed transformation applied to the gradient image of a filtered version of the channel f_c (see also Fig. 7.22).

[3]The gray-level saturation S_α should not be confounded with the color saturation.

First, we attenuate the noise in the image using a Gaussian filter G (type and parameters of the filter are not crucial, we used a 9×9 filter with $\sigma = 4$). Then, the vessels interrupting the circular shape of the optic disk are filled using a morphological closing:

$$p_1 = \phi^{(s_1 B)}(G * f_c) \tag{7.25}$$

with s_1 such that the largest vessels are filled (as explained in the previous section). In order to remove irregularities within the papillary regions that may also produce a high-gradient value, we apply an opening by reconstruction:

$$p_2 = R_{p_1}(\varepsilon^{(s_2 B)}(p_1)) \tag{7.26}$$

$s_2 = 15$ has been found to be a good value for 640×480 images. This is a big opening, but thanks to the reconstruction, the contours of p_1 are preserved.

Then, the morphological gradient is calculated:

$$\rho p_2 = \delta^{(B)} p_2 - \varepsilon^{(B)} p_2 \tag{7.27}$$

Calculating the watershed transformation of this gradient would lead to a strongly oversegmented result. Once again, we have to find a marker and impose it (see section 7.3). With only one source within the optic disk, the algorithm gives exactly one catchment basin which—if the filtering process has been efficient—coincides exactly with the optic disk. We use the approximated center c as "inner marker." As external marker, we use a circle centered in c with a diameter larger than two times the largest possible diameter of the papilla (factor 2 for the case that the approximation was bad and the approximation of the center c lies on the border of the optic disk).

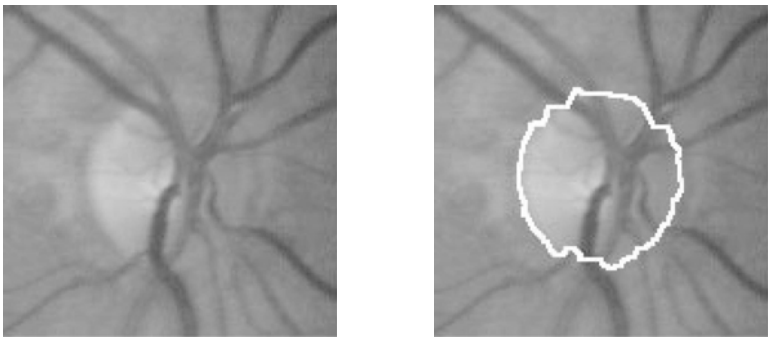
$$m(x) = \begin{cases} \rho p_2 & \text{if} \quad x \in \{c\} \cup Cercle(c) \\ t_{\max} & \text{if} \quad x \notin \{c\} \cup Cercle(c) \end{cases} \tag{7.28}$$

With this marker, we can now calculate the watershed transformation:

$$P_{fin} = CV_i \left[R^*_{\rho p_2}(m) \right] \quad \text{with } c \in CV_i \tag{7.29}$$

7.5.2.5 Results

The algorithm has been tested on 60 color fundus photographs (640×480) taken with a Sony color video 3CCD camera on a Topcon TRC 50 IA



(a) The optic disk in a color image

(b) Segmentation result

Figure 7.23: Detection of the optic disk.

retinograph. These images have not been used for the development of the algorithm.

The optic disk has been localized correctly in 57 of these 60 images. In 3 of these 60 images, there were very large accumulations of exudates which inhibited a correct localization of the optic disk. The accuracy of the detection of the contours has been assessed qualitatively by a human grader; there were 48 images, for which the segmentation result was satisfying, with no or few pixels missed or falsely detected (e.g. see Fig. 7.23). In eight images, there were some parts missing due to very poor contrast of the original images, but the result contained still more than 75% of the optic disk. In one image, the result was not satisfying, once again due to low contrast: Indeed the contour was hardly visible, even for a human.

7.6 The Detection of Pathologies in Color Fundus Images

Pathology detection is certainly the most important part of analysis of retinal images. In diabetic retinopathy, there are three types of lesions indicating different stages of the disease that can be detected using color fundus images: microaneurysms, exudates, and hemorrhages. In this section, we present automatic algorithms for the detection of microaneurysms and exudates. An algorithm for the detection of hemorrhages can be found in [9].

7.6.1 The Detection of Microaneurysms

7.6.1.1 Motivation

Microaneurysms are the first ophthalmoscopic sign of diabetic retinopathy [1]. Over and above that, their number is an indication of the progression of the disease. Their detection is therefore crucial for the diagnosis of diabetic retinopathy, for the mass screening, *and* for the monitoring of the disease.

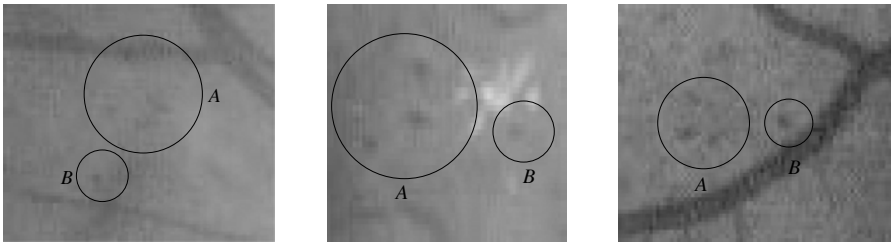
7.6.1.2 Properties

Microaneurysms are tiny dilations of the capillaries. They appear as small reddish isolated patterns of circular shape in color fundus images of the human retina [1]. Their diameter normally lies between 10 and 100 μm, but it is always smaller than 125 μm. As they come from capillaries, and as capillaries are not visible in color fundus images, they appear as isolated patterns, i.e. disconnected from the vascular tree.

Microaneurysms are sometimes hard to detect: Their contrast is often very low and sometimes, they are hardly visible and difficult to distinguish from noise. Their reddish color can hardly be used for their detection, because it is far from being constant in different images (see Fig. 7.24).

7.6.1.3 State of the Art

The first algorithm for the detection of microaneurysms has been presented Laÿ [19]. The author introduced the radial opening $\gamma^{\sup} = \bigcup \gamma^{L_i}$, i.e. the supremum



(a) (b) (c)

Figure 7.24: Microaneurysms in color images. (a) Sure microaneurysms; (b) doubtful cases.

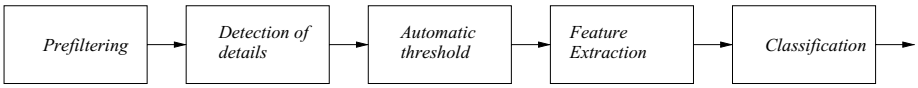


Figure 7.25: The principle of the algorithm for microaneurysm detection.

of openings with linear structuring elements L_i in different directions in order to remove the microaneurysms but to preserve the piecewise linear vessels. This technique has been used by nearly all the authors working on the automatic detection of microaneurysms; important improvements have been proposed in [20, 21].

7.6.1.4 The Algorithm

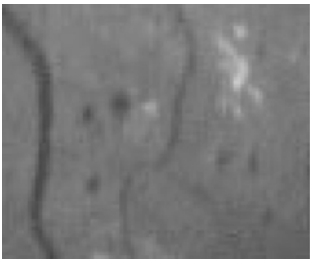
The algorithm presented in this section is based on the strategy shown in Fig. 7.25. First, the shade correction method described in section 7.4 is applied, and then candidates are detected by means of the diameter closing and an automatic threshold; features calculated for these candidates allow their classification into real microaneurysms and false positives. A first version of this algorithm has been presented in [22].

Prefiltering and shade correction: The objective of this step is to attenuate the noise, to enhance the contrast, and to correct the nonuniform illumination.

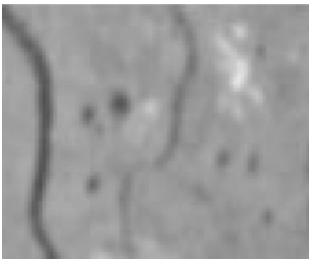
As it has been stated in section 7.2, microaneurysms—like all blood containing elements—appear best contrasted in the green channel. First, the shade correction operator described in section 7.4 is applied to the green channel. It is crucial that this algorithm does not introduce new dark regions which would cause a lot of false positives.

Besides the shade correction, the operator $\mathrm{SC}_{\mathrm{norm}}$ enhances the contrast of structures in the image depending on their size. In order to privilege small vessels and microaneurysms more than larger hemorrhages and large vessels, an adapted size of the window used in $\mathrm{SC}_{\mathrm{norm}}$ can be chosen. A small gaussian filter attenuates the noise, but enhances microaneurysms; it can be seen as a matched filter [20]. With G a gaussian filter, we obtain the prefiltered image by (see also Fig. 7.26):

$$p = G * \mathrm{SC}_{\mathrm{norm}}(f_{\mathrm{g}}) \tag{7.30}$$



(a) A detail of a fundus image containing microaneurysms



(b) Detail of the prefiltered image

Figure 7.26: Prefiltering step.

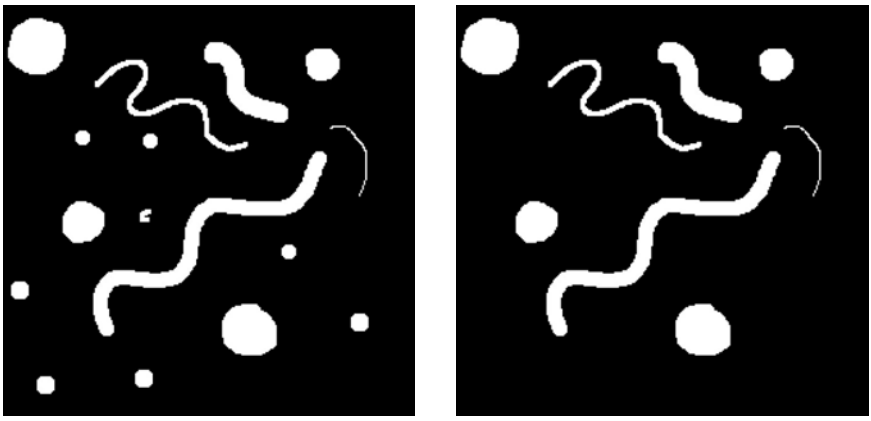
The detection of dark isolated details by means of the diameter closing: The next step is to find the "candidates," i.e., all features that may possibly correspond to microaneurysms. Microaneurysms are characterized by their diameter; in the green channel of a color image, they correspond to dark details—"holes"—with a maximal diameter of λ (with λ depending on the image resolution).

As in the top-hat transformation used for vessel detection in section 7.5.1, the main idea is to first construct a closing ϕ that removes the details from the image and then calculate the difference to the original image. However, a morphological closing cannot be used in our case because it fills not only the holes but also the ditches (vessels). One possibility to fill only the holes without filling the ditches is to determine the infimum of openings with linear structuring elements in different directions, because they do fit into the vessels in at least one direction. However, this is only an approximative solution of the problem; a tortuous line for example will be closed as well. We will now present the diameter closing $\phi_\lambda^\diamond$ which removes all dark details of a diameter smaller than λ.

First, we define the diameter α of a connected set X as its maximal extension, i.e. the maximal distance between two points of the set:

$$\alpha(X) = \bigvee_{x,y \in X} d(x,y) \tag{7.31}$$

with $d(x, y)$ the distance between two points x and y. For simplicity, we use the block distance: If $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{Z}^2$ are two points and x_1, x_2 and y_1, y_2 their coordinates, respectively, the block distance can be written as $d(x, y) = |x_1 - y_1| \vee |x_2 - y_2|$.



(a) A binary image (b) The result of a diameter opening

Figure 7.27: The diameter opening of a binary image: all connected components with a diameter inferior to 15 pixels are removed.

With this definition of the diameter of a set, we can define a trivial opening. Let X be an arbitrary binary image and X_i its connected components, i.e. $X = \bigcup X_i$ and $X_i \cap X_j$ for $i \neq j$. The diameter opening is the union of all connected components X_i with a diameter greater or equal to λ (see Fig. 7.27):

$$\gamma_\lambda^\circ(X) = \bigcup_{\alpha(X_i)\geq\lambda} X_i \tag{7.32}$$

As the applied criterion $\alpha(X_i) \geq \lambda$ is increasing, i.e., $X \subset Y$ implies that if X fulfills the criterion, Y also does, the operation $\gamma_\lambda^\circ(X)$ is an opening.

It can be shown that the diameter opening is the supremum of all openings with structuring elements with a diameter greater than or equal to λ [8]:

$$\gamma_\lambda^\circ(X) = \bigcup_{\alpha(B)\geq\lambda} \gamma^B(X) \tag{7.33}$$

It is, therefore, a generalization of the approximative method proposed in [16] used by the majority of authors, where only linear structuring elements fulfilling the criterion are used.

The diameter closing removes all holes X_i^c (connected components of the background X^c) with a diameter inferior λ. Furthermore, it can be written as the infimum of all morphological closings with structuring elements whose diameter

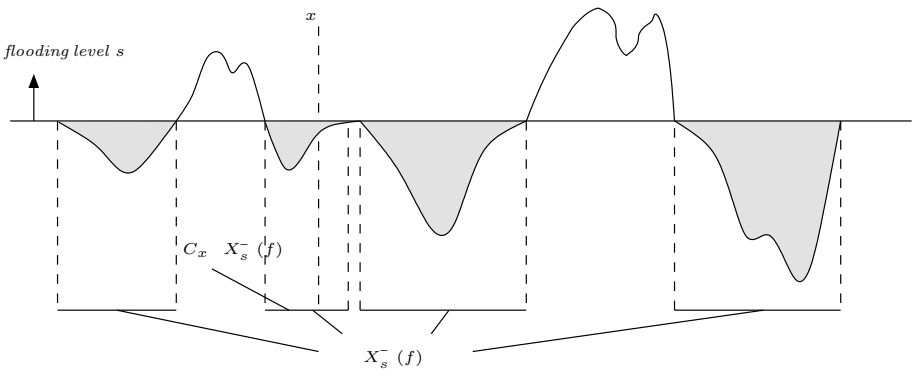


Figure 7.28: The flooding of an image f at level s.

is equal or superior to λ:

$$\begin{aligned}\left[\phi_\lambda^\circ(X)\right](x) &= X \cup \left(\bigcup_{\alpha(X_i^c)<\lambda} X_i^c \right) \\ &= \bigcap_{\alpha(B)\geq\lambda} \phi^B \end{aligned} \tag{7.34}$$

We have now defined the diameter opening and closing for the binary case. In order to pass from binary to gray-level images, we can apply the binary operator to all level sets (the results of threshold operations for all gray levels $t \in T$). Let $C_x(X)$ be the connected opening, i.e., the connected component of X containing x if $x \in X$ and the empty set if $x \notin X$. Furthermore, let $X_t^+(f)$ be the section of f at level t, i.e., the set of all pixels for which $f(x) \geq t$ and $X_t^-(f)$ the section of the background (the "lakes," see Fig. 7.28):

$$\begin{aligned} X_t^+(f) &= T_{[t,t_{\max}]}(f) = \{x \in E \mid f(x) \geq t\} \\ X_t^-(f) &= T_{[t_{\min},t]}(f) = \{x \in E \mid f(x) \leq t\} \end{aligned} \tag{7.35}$$

Then, the gray scale diameter opening and closing can be defined respectively:

$$\begin{aligned} \gamma_\lambda^\circ(f) &= \sup\left\{s \leq f(x) \mid \alpha\left(C_x\left[X_s^+(f)\right]\right) \geq \lambda\right\} \\ \phi_\lambda^\circ(f) &= \inf\left\{s \geq f(x) \mid \alpha\left(C_x\left[X_s^-(f)\right]\right) \geq \lambda\right\} \end{aligned} \tag{7.36}$$

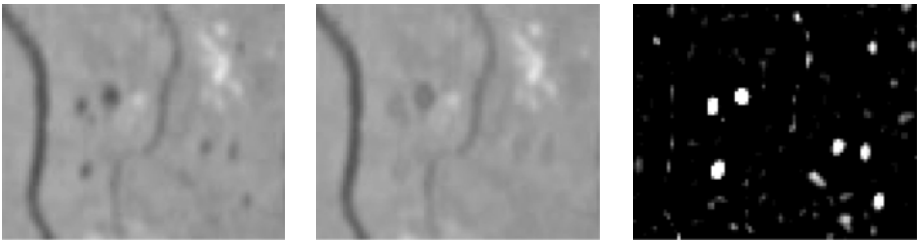
Of course, Eq. (7.36) cannot be used for implementation of this algorithm because it would be highly inefficient. Instead of calculating the diameter opening

for each threshold, we use hierarchical queues in order to simulate a flooding of the image. We explain this technique for the diameter closing.

The flooding starts with the lowest local minima in the image (i.e. with the global minima). We determine the diameter of all the lakes with gray level s. If the diameter of a lake exceeds λ, the output image takes the value s for all the points belonging to this lake ("the flooding stops for this lake"). Then s is incremented, the new local minima at this level are added, and the existing lakes are extended until there is no more pixel x in the image with $f(x) \leq s$ not belonging to a lake. If two lakes meet, they fuse and it is considered as one lake from now on. Then, when the flooding has been finished for this level, the diameter of all lakes are calculated. If the diameter of a lake exceeds λ, but has not exceeded λ for the previous level, the output image is set to s for all the pixels of this lake. In this way, we flood the whole image until there is no more lake with a diameter inferior to λ.

This algorithm can be implemented very efficiently with hierarchical queues. See [8] for details.

In Fig. 7.29, we show the application of the diameter closing to the detection of microaneurysms. The prefiltered image is shown in Fig. 7.29(a); its closing by diameter in Fig. 7.29(b). We note that the distinction of holes and ditches (microaneurysms and vessels) works very well: The microaneurysms are completely filled whereas the vessels are not touched. The associated top-hat $\phi_\lambda^\circ f - f$ is shown in Fig. 7.29(c). The small details visible in this image and not corresponding to microaneurysms are "parasite holes" and are due to irregularities and noise in the image. From this image, it is easy to get the candidates by a threshold. The applied threshold technique is shown in the next paragraph.



(a) The prfiltered and shade corrected image

(b) The diameter closing

(c) The associated top-hat transformation

Figure 7.29: Detection of dark details by means of the diameter closing.

The automatic threshold: The threshold can be seen as the minimal contrast a detail must have in order to be considered as a candidate.

If the threshold is chosen manually, we lose the main advantages of an entirely automatic analysis. If a fix threshold is applied, we have to deal with a lot of false positives or with poor sensitivity, because the contrast of microaneurysms may be very different from one image to another. If it depends exclusively on the histogram of the top-hat image, it is supposed that the image contains microaneurysms. Hence, we have to find a compromise between a fix a histogram-dependent threshold.

In order to find an automatic method for the determination of an automatic threshold, we have analyzed 10 retinal images. For all these images, we have chosen an "optimal" threshold using ROC-analysis, i.e., a threshold that gives the best compromise between sensitivity and number of false positives.

This optimal threshold has then been compared to statistical properties of the top-hat image (standard deviation, amount of noise, volume of the top-hat image, etc.). The most obvious relation has been found between the volume of the top-hat image and the optimal threshold. This relation is shown in Fig. 7.30.

This result is not really surprising. The volume of the top-hat image depends on two image properties: the contrast and the amount of noise. On the one hand,

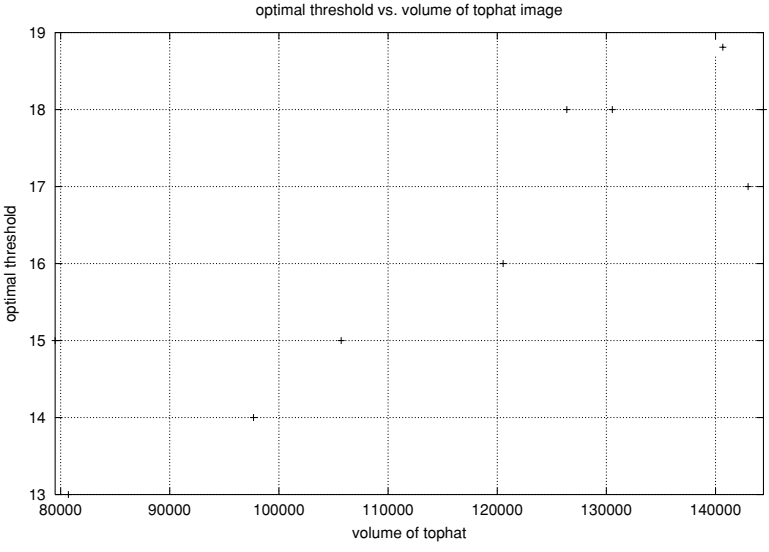


Figure 7.30: Optimal threshold versus volume of top hat.

the better the contrast is, the higher the threshold *can* be chosen. On the other hand, the higher the amount of noise is, the higher the threshold *must* be chosen.

However, some "fix" information must be incorporated by using lower and upper bounds for the threshold:

$$t_{vol}(V) = \begin{cases} 13 & \text{if } V < 80000 \\ 10^{-4} \cdot V + 5 & \text{if } 80000 \leq V \leq 130000 \\ 18 & \text{if } V > 130000 \end{cases} \tag{7.37}$$

The candidate regions are determined by a double threshold technique (see [6] for details). This technique allows one to apply a lower threshold without accepting a higher number of candidates:

$$\begin{aligned} CA_1 &= T_{[t_{\text{vol}}, t_{\max}]}\left(\vartheta_{\phi_\lambda^\circ}(f)\right) \\ CA_2 &= T_{[\frac{2}{3} \cdot t_{\text{vol}}, t_{\max}]}\left(\vartheta_{\phi_\lambda^\circ}(f)\right) \\ CA &= R_{CA_2}(CA_1) \end{aligned} \tag{7.38}$$

This improvement in the determination of the candidate region is important, because the features that are calculated for the candidates depend a lot on this region, as we will see in the following paragraph.

Elimination of the candidates situated on the vessels: Before calculating the features, we can exclude all candidates situated on the vascular tree. As we have seen, a top-hat transformation associated to a morphological closing extracts all dark details that cannot contain the structuring element, i.e., all "holes" and all "ditches," and as a consequence all microaneurysms and all vessels. Comparing the morphological top-hat transformation with the one associated to the diameter closing of the same size, we can identify the false candidates situated on vessels and hemorrhages: For candidates not situated on vessels, we can assume that the values of the two top-hat images are approximately the same:

$$\left[\vartheta_{\phi^{(sB)}}(p)\right](x) \approx \left[\vartheta_{\phi_\lambda^\circ}(p)\right](x) \tag{7.39}$$

This is not the case for candidates situated on the vessels. We can write the modified candidate image CA' as

$$CA' = \left\{x \in CA \mid \left[\vartheta_{\phi^{(sB)}}(p)\right](x) \leq 2 \cdot \left[\vartheta_{\phi_\lambda^\circ}(p)\right](x)\right\} \tag{7.40}$$

Candidates situated on the optic disk can be easily removed using the segmentation result from section 7.5.2.

Feature extraction and classification: With the top-hat transformation and the automatic threshold, we have found candidates, i.e. possible microaneurysms, using just a size criterion and a contrast measure (threshold). However, there are still many false positives, and the result is not acceptable. But there are still other properties to be exploited. We used the following features in order to classify the candidates into true microaneurysms and false positives:

- *The surface:* Fundus images are often corrupted by noise (high frequency gray level variations). Hence, there are many small "holes" and "peaks" in the image; therefore, the surface of the candidate regions is an important feature:

$$\mathrm{Surf}(C_i) = \#C_i \tag{7.41}$$

- *The circularity:* We have used the maximal extension as a feature. That means that small linear features are also extracted. The circularity may help excluding them:

$$\mathrm{Circ}(C_i) = \frac{\mathrm{Surf}(C_i)}{(\alpha(C_i))^2} \tag{7.42}$$

- *The maximal value of the top-hat image:* In the threshold operation, we have already used this feature. On the other hand, it may be important to combine it with other features.

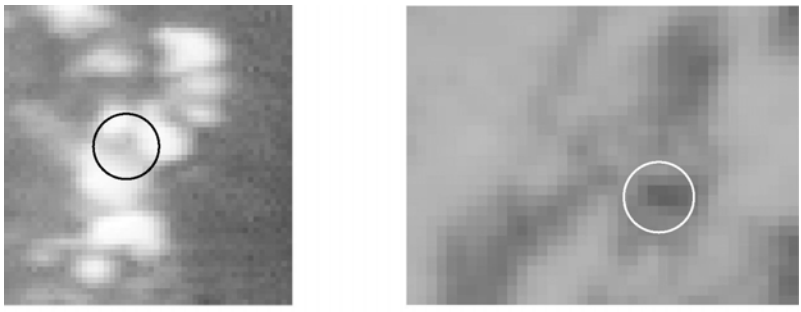
$$MV_{\vartheta_{\phi_\lambda^\circ} p}(C_i) = \max_{x \in C_i} \{\vartheta_{\phi_\lambda^\circ} p(x)\} \tag{7.43}$$

- *The dynamic:* The dynamic is a measure of "deepness" of a minimum. If a minimum is very deep or in contrary very shallow, it is probably not a microaneurysm.

- *The outer mean value:* It is also important to take into consideration the absolute gray level values on the outside of the candidate. The mean on the external gradient can help finding false positives due to exudates or hemorrhages (see Fig. 7.31):

$$\begin{aligned} Ex(C_i) &= \delta^{3B} C_i \setminus \delta^B C_i \\ \mu_{\mathrm{ext}} &= \frac{1}{\#Ex(C_i)} \sum_{x \in Ex(C_i)} p(x) \end{aligned} \tag{7.44}$$

- *The contrast measure:* The maximal value of the top-hat image is a contrast measure: It is the difference between the local minimum and the level for which the flooding stops. Another contrast measure is the difference



(a) An exudate causing a false positive

(b) An hemorrhage causing a false positive

Figure 7.31: Two types of false positives that can be identified using the mean value of the prefiltered image on the external gradient of the candidate.

between the mean value on the external gradient of the candidate region and the mean value on the candidate region itself:

$$\begin{aligned} \mu_{\text{int}}(f) &= \frac{1}{\#C_i} \sum_{x \in C_i} f(x) \\ \mu_{\text{ext}}(f) &= \frac{1}{\#Ex(C_i)} \sum_{x \in Ex(C_i)} f(x) \\ \text{contr}_f(C_i) &= \mu_{\text{ext}}(f) - \mu_{\text{int}}(f) \end{aligned} \tag{7.45}$$

- *The color:* We have already seen in the section 5.2 that the green channel contains the most important information about blood-containing elements in the retina and this is why it is used for the detection of microaneurysms. However, there is also some information in the red, and sometimes in the blue channel. We have studied a lot of color features; the most efficient are the following two:

  1. *Color Contrast in the Luv color space:* In the Luv color space, the euclidean distance can be seen as the "true" distance, i.e. the perceptible distance. We used, therefore, the euclidean distance between the color on the candidate region and the color on its external gradient:

$$\begin{aligned} \text{contr}_{Luv}(C_i) = \big[&(\mu_{\text{ext}}(L) - \mu_{\text{int}}(L))^2 + (\mu_{\text{ext}}(u) - \mu_{\text{int}}(u))^2 \\ &+ (\mu_{\text{ext}}(v) - \mu_{\text{int}}(v))^2\big]^{\frac{1}{2}} \end{aligned} \tag{7.46}$$

  2. *Contrast of the principal components of the red and the blue channel:* In order to find color information complementary to the

information in the green channel, we use the principal component cp_{rb} of the blue and the green channel as a feature:

$$\mathrm{contr}_{cp_{rb}}(C_i) = \mu_{\mathrm{ext}}(cp_{rb}) - \mu_{\mathrm{int}}(cp_{rb}) \tag{7.47}$$

These two features do not depend strongly on each other. They help identifying some false positives, but their efficiency is limited.

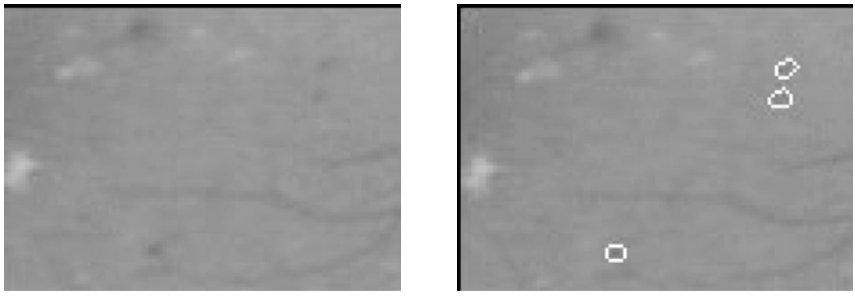
For the classification, a KNN-classifier is used (K-nearest neighbors), for it has been shown to work well even if there are outliers [23, 24]. We do not detail this method, for it is a standard method of classification.

As training set, we used a set of 16 images. We asked two ophthalmologists to mark the microaneurysms independently and then to compare and discuss their results. They finally agreed on 201 microaneurysms; this has been taken as a *golden standard*. Our algorithm was then applied on these images; 924 candidates were found. Among them were 199 true positives. These candidates were used to train the classifier.

7.6.1.5 Results

The algorithm has been tested on 57 images and the results have been compared to the ones obtained by two human graders: As for the training set, the specialists graded the images independently, then they compared and discussed the results. The result of this procedure was considered as *golden standard.*

The comparison with the automatic method gave a mean sensitivity of 88.1% and a predictive value of 83.8% (2.3 FP per image). In Fig. 7.32, an example is shown.



(a) Microaneurysms in a color image (b) Detected microaneurysms

Figure 7.32: Result of microaneurysm detection in color images.

7.6.2 The Detection of Hard Exudates

7.6.2.1 Motivation

Hard exudates are yellowish intraretinal deposits made up of serum lipoproteins. They are the result of leaking from the abnormally permeable blood vessels, especially microaneurysms.

Hard exudates may be observed in several retinal vascular pathologies, but are a main hallmark of diabetic macular edema. If macular edema is diagnosed in an early and still asymptomatic stage, laser treatment can be very efficient and prevent vision loss. In a screening context, the easiest way of detecting macular edema is to detect hard exudates and their distance to the macula.

7.6.2.2 Properties

Exudates appear as bright patterns in color fundus images [1]. They are characterized by a strong contrast; their shape and size are completely variable, and their contours mostly irregular.

However, they are not the only bright features in retinal images; the optic disk and eventual over-exposed regions have similar gray levels. Regions surrounded by vessels may also be bright and well contrasted.

7.6.2.3 State of the Art

In [25], the authors propose shade correction and image enhancement techniques. Then, a threshold is manually chosen in order to detect the exudates. We think that a full automation of exudate detection is possible and useful.

In [26], a method based on image enhancement, shade correction, and a combination of local and global thresholding is proposed and validated.

The method proposed in [28] is based on shade correction and advanced classification methods.

7.6.2.4 The Algorithm

Our algorithm can be subdivided into two parts: First, we find candidate regions, i.e. regions that possibly contain exudates. In a second step, we determine the

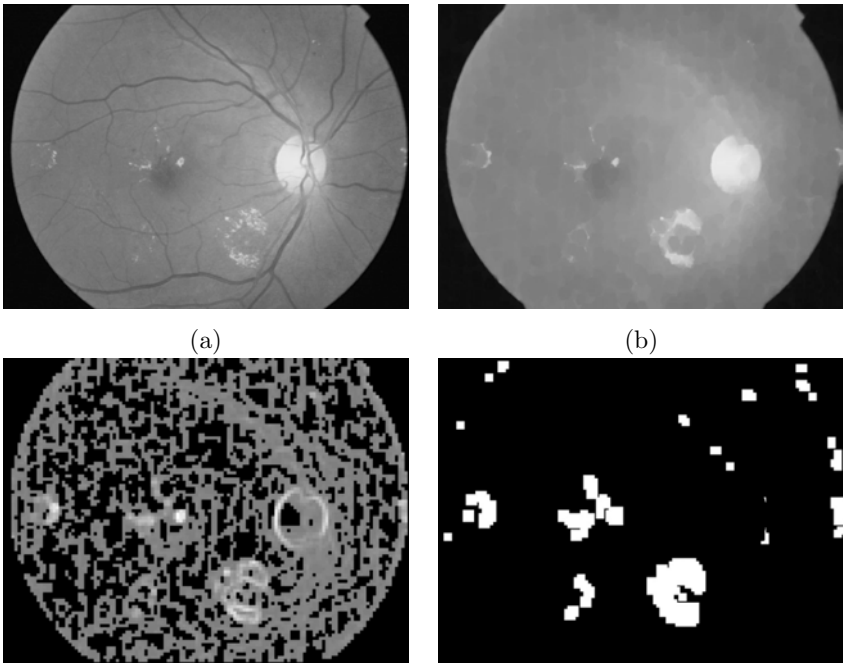


Figure 7.33: (a) The luminance channel of a color image of the human retina. (b) The closing of the luminance channel. (c) The local standard variation in a sliding window. (d) Candidate region.

contours of the exudates. This algorithm has been published and discussed in [18]; here in we give a sketch of it.

Finding the candidate regions: Regions containing exudates are characterized by a high contrast *and* a high gray-level. The problem that occurs, if we use the local contrast to determine regions that contain exudates, is that bright regions surrounded by dark vessels may also produce a high local contrast. As shown in the section 7.3, vessels can be removed by means of a morphological closing (see Fig. 7.33(b)):

$$e_1 = \phi^{(s_1 B)}(f_g) \tag{7.48}$$

On this image we calculate the local variation for each pixel x within a window $W(x)$ (see Fig. 7.33(c)) centered in x:

$$e_2(x) = \frac{1}{N-1} \cdot \sum_{\xi \in W(x)} (e_1(\xi) - \mu_{e_1}(x))^2 \tag{7.49}$$

In order to spare computational time, e_2 is not calculated for every pixel; it is calculated for a subsampled version of e_1. Then e_2 is found by interpolation.

Applying a fix threshold on the image e_2 at gray level α_1, we obtain all regions with a standard variation larger than or equal to α_1. However, bright objects larger than the window do produce only a high standard variation on its borders. In order to obtain the whole candidate regions, we fill the holes by reconstructing the image from its borders Bo_f [6]. We also dilate the candidate region in order to ensure that there are background pixels next to exudates that are included in the candidate regions:

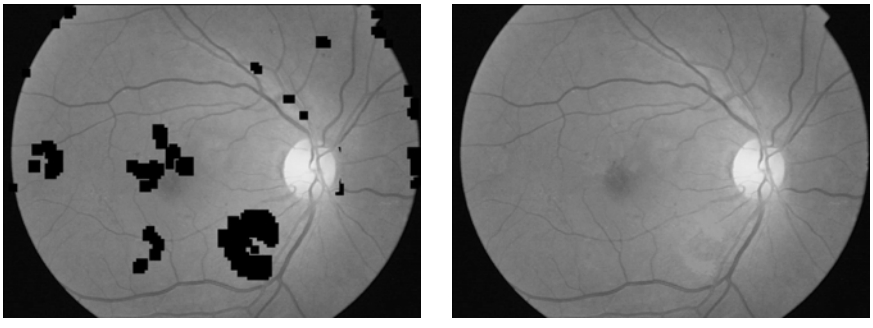
$$\begin{aligned} e_3 &= \delta^{(sB)}\left(T_{[\alpha_1, t_{\max}]}(e_2)\right) \\ e_4 &= R^*_{e_3}(b) \quad \text{with } b = \begin{cases} 0 & \text{if } x \in Bo_f \\ t_{\max} & \text{if } x \notin Bo_f \end{cases} \end{aligned} \tag{7.50}$$

The threshold α_1 is chosen favoring sensitivity to specificity: False positives can be identified later. Then, we remove a dilated version of the optic disk and we obtain the candidate regions:

$$ca = e_4 - e_4 \wedge \delta^{(sB)}(p_{fin}) \tag{7.51}$$

Finding the contours: In order to find the contours of the exudates, we set all the candidate regions to 0 in the original image (see Fig. 7.34(a)):

$$m(x) = \begin{cases} 0 & \text{if } ca(x) \neq 0 \\ f_{\mathrm{g}}(x) & \text{if } ca(x) = 0 \end{cases} \tag{7.52}$$



(a) (b)

Figure 7.34: (a) The candidate regions set to 0 in the original image. (b) The morphological reconstruction.

and then we calculate the morphological reconstruction by dilation of the resulting image under f_g (see Fig. 7.34(b)). Exudates are now completely removed from the image, as they are completely comprised in the candidate regions. We can, therefore, calculate the difference to the original image and apply a fix threshold in order to obtain the final segmentation result:

$$e_{\text{fin}} = T_{[\alpha_2, t_{\max}]}(f_g - R_{f_g}(m)) \tag{7.53}$$

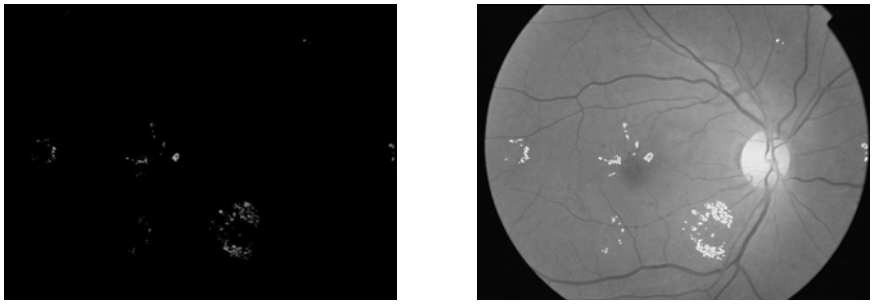
This algorithm has three parameters: The size of the window W and the two thresholds α_1 and α_2. The choice of the size of W is not crucial, and we have found good results for a window size of 10×10. If the window size is very large, small isolated exudates are not detected. From a medical point of view, this is not really problematic. The first threshold α_1 determines the minimal variation value within the window that is suspected to be a result of the presence of exudates. If α_1 is chosen too low, the number of false positives increases, if it is set too high, sensitivity decreases. The parameter α_2 is a contrast parameter: It determines the minimal value a candidate must differ from its surrounding background to be classified as an exudate.

7.6.2.5 Results

We have tested the algorithm on an image data base of 30 digital images 640×480 taken with a Sony color video 3CCD camera on a Topcon TRC 50 IA retinograph. These images have not been used for the development of the algorithm. Fifteen of these images did not contain exudates, and in 13 of these 15 no exudates were found by our algorithm. In two images, few false positives were found (less than 20 pixels).

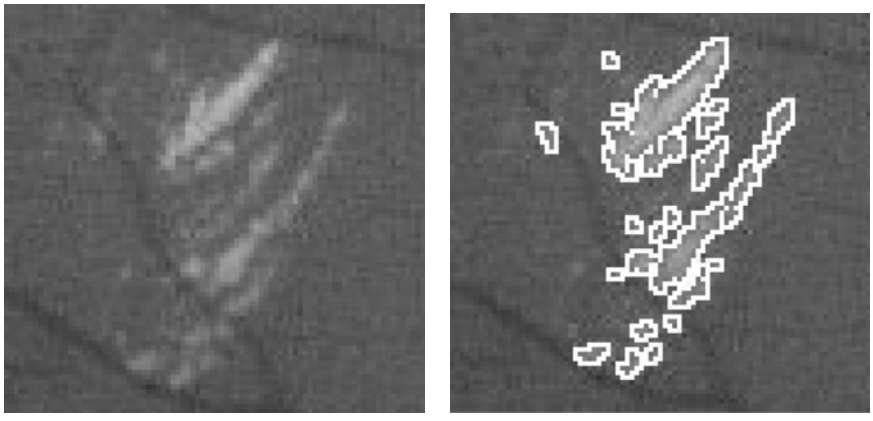
We asked an ophthalmologist to mark the exudates in the 15 images and compared the results obtained by the algorithm to his. The comparison was done pixel-wise (with 1 pixel tolerance), and as for exudates, the number cannot be determined; it is the surface and the position rather than the number which can be used for diagnostic purposes.

We obtained a mean sensitivity of 92.8% and a predictive value of 92.4%. In Fig. 7.35, an example for the automatic detection of exudates is shown (see also Figs. 7.36 and 7.37).



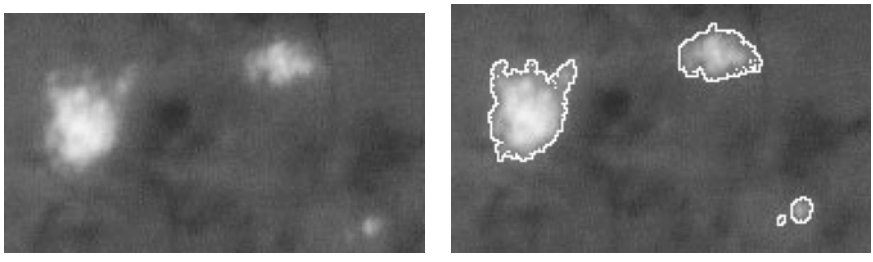
(a) The top-hat image (b) Algorithm result

Figure 7.35: The result of exudates detection.



(a) (b)

Figure 7.36: (a) A detail of the green channel of a color image containing exudates. (b) The segmentation result.



(a) (b)

Figure 7.37: (a) A detail of the green channel of a color image containing exudates. (b) The segmentation result.

7.7 Conclusion and Perspectives

In this chapter, we have seen different ways of computer assistance to the diagnosis of diabetic retinopathy, which is a very frequent and severe eye-disease: image enhancement, mass screening, and monitoring. Different algorithms within this framework have been presented and evaluated with encouraging results.

However, there are still improvements to be made. The first one is to use high-resolution images. We worked on images already used in centers of ophthalmology, but it is clear that acquisition techniques also improve and that in the coming years high-resolution images will become clinical standard. Future segmentation algorithm can make use of this high resolution (e.g. there will be more features for microaneurysm detection).

Another possible research axis is the inclusion of patient data into the algorithms. This *a priori* knowledge about the patient is used by physicians; it also could be used by automatic methods. For instance, we have observed, that the color of black people's eyes is quite different from the color of white people's, the color of a child's retina is different from the color of an adult's eye. This is precious information that could be used in order to enhance the performance of lesions detection algorithms.

Even if there is still progress to be made, the presented algorithms work well; a clinical trial is envisaged.

7.8 Annex: Algorithm Evaluation

Whenever objects are detected automatically, the performance of the algorithm has to be evaluated. In the medical domain, results are normally compared to the results obtained by one or more specialists.

Let us consider a medical examination (diagnostic test). Often, such a test can only be positive or negative (the patient suffers from the disease or not). In order to evaluate the efficiency of this diagnostic test, its result is compared to reality; "the truth" is found by other diagnostic methods. For this comparison, we define:

- *True Positive (TP):* The patient suffers from the disease and the test was positive.

- *False Positive (FP):* The patient does not suffer from the disease, but the test was positive.
- *True Negative (TN):* The patient does not suffer from the disease, and the test was negative.
- *False Negative (FN):* The patient suffers from the disease, but the test was negative.

With these definitions, we can evaluate the performance of a diagnostic test by means of sensitivity and specificity, defined as

$$\begin{aligned} \text{sensitivity} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}} \end{aligned} \tag{7.54}$$

$\text{TP} + \text{FN}$ is the number of patients suffering from the disease and $\text{TN} + \text{FP}$ is the number of patients not suffering from the disease; the sensitivity is the percentage of detected cases of the disease and the specificity is the percentage of correctly classified healthy persons.

These definitions can be transfered to the evaluation of detection/classification algorithms, i.e. true positives are correctly detected pathologies, false positives are nonpathological objects falsely classified by the algorithm, etc.

There is, however, a difference between detection and classification algorithms: in detection problems, the number of objects is not limited as it is the case for classification problems (e.g. the classification of patients). In detection problems, a definition of true negatives does not make sense. There are two possibilities to resolve this problem:

- If the number of objects is an important quantity (number of lesions, e.g. microaneurysms), then the number of false positives may be a good indicator for the quality of the algorithm.
- If the number cannot be determined or if it is not the important quantity (this is the case if these are strong variations in shape and size of lesions—like for exudates for example), a pixel-wise comparison between the two results is preferable. In this case, the predictive value can be calculated:

$$\text{pv} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7.55}$$

This is the probability that an object (or pixel) classified as positive is really positive.

With these values (sensitivity, number of false positives, predictive value), the quality of automatic pathology detection algorithms can be assessed.

7.9 Acknowledgment

First of all, the authors thank the ophthalmology department of the Lariboisire Hospital in Paris for their excellent collaboration, their hearty and competent support, for having supplied all images, and for having evaluated the performance of all algorithms presented in this chapter.

This work has been supported by the French Ministry of Education and Research (MENRT) in the program Dpistage automatique de la rétinopathie diabtique (00 B 0100 01).

Questions

1. *Why do usual shade correction algorithms darken pixels close to bright objects?*
2. *How can this darkening effect be prevented?*
3. *What is the difference between an algebraic and a morphological closing?*
4. *How can dark details in a gray scale image be extracted using mathematical morphology?*
5. *How can dark details with a maximal extension of λ be extracted?*
6. *How does the use of markers in the watershed transformation work and what is their influence on the result?*
7. *How can the watershed transformation be used for the detection of thin dark lines in a gray scale image?*
8. *How can the watershed transformation be used for the detection of object contours?*

9. *Which morphological operator can be used for removing bright objects preserving the borders of all remaining objects?*

10. *In the analysis of fundus images, specificity cannot be used for an assessment of the quality of a pathology detection algorithm. Why?*

Bibliography

[1] Massin, P., Erginay, A., and Gaudric, A., Rétinopathie Diabétique, Elsevier, Editions scientifiques of médicales, Elsevier, SAS, Paris, 2000.

[2] Lee, S. C., Lee, E. T., Kingsley, R. M., Wang, Y., Russell, D., Klein, R., and Warn, A., Comparison of diagnosis of early retinal lesions of diabetic retinopathy between a computer system and human experts, Arch. Ophthalmol., Vol. 119, pp. 509–515, 2001.

[3] Delori, F. C. and Pflibsen, K. P., Spectral Reflectance of the Ocular Fundus, Appl. Optics, Vol. 28, pp. 1061–1071, 1989.

[4] Preece, S. J. and Claridge E., Monte Carlo modelling of the spectral reflectance of the human eye, Phy. Med. Biol., Vol. 47, pp. 2863–2877, 2001.

[5] Serra, J., Image Analysis and Mathematical Morphology, Academic Press, San Diego, CA, 1988.

[6] Soille, P., Morphological Image Analysis: Principles and Applications, Springer-Verlag, Berlin, 1999.

[7] Beucher, S. and Meyer, F., The morphological approach to image segmentation: The watershed transformation, In: Mathematical Morphology in Image Processing, Dougherty, E. R., ed., Marcel Dekker, New York, pp. 433–481, 1992.

[8] Vincent, L., Morphological area openings and closings for grayscale images, In: NATO Shape in Picture Workshop, Driebergen, 1992, pp. 197–208.

[9] Walter, T., Application de la Morphologie Mathématique au diagnostic de la Rétinopathie Diabétique à partir d'images couleur, Ph.d. Thesis, Centre of Mathematical Morphology, Paris School of Mines, September 2003.

[10] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., and Goldbaum, M., Detection of blood vessels in retinal images using two-dimensional matched filters, IEEE Trans. Med. Imaging, Vol. 8, No. 3, pp. 263–269, 1989.

[11] Zana, F. and Klein, J.-C., Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation, IEEE Trans. Image Process., Vol. 10, No. 7, pp. 1010–1019, 2001.

[12] Sahoo, P. K., Soltani, S., Wong, A. K. C., and Chen, Y. C., A survey of Thresholding Techniques, Comput. Vision, Graphics, Image Process., Vol. 41, pp. 233–260, 1988.

[13] Sinthanayothin, C., Boyce, J. F., Cook, H. L., and Williamson, T. H., Automated localisation of the optic disc, fovea and retinal blood vessels from digital colour fundus images, Br. J. Ophthalmol., Vol. 83, No. 8, pp. 231–238, 1999.

[14] Tamura, S. and Okamoto, Y., Zero-crossing interval correction in tracing eye-fundus blood vessels, Patt. Recogn., Vol. 21, No. 3, pp. 227–233, 1988.

[15] Pinz A., Prantl, M., and Datlinger P., Mapping the human retina, IEEE Trans. Med. Imaging, Vol. 1, No. 1, pp. 210–215, 1998.

[16] Osareh, A., Mirmehdi, M., Thomas, B., and Markham, R., Colour morphology and snakes for optic disc localisation, In: The 6th Medical Image Understanding and Analysis Conference, 2002, pp. 21–24.

[17] Walter, T. and Klein, J.-C., Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques, In: Lecture Notes in Computer Science, Vol. 2199, Crespo, J., Maojo, V., and Martin, F., eds., Springer-Verlag, Berlin, pp. 282–287, 2001.

[18] Walter, T. and Klein, J.-C., A contribution of image processing to the diagnosis of diabetic retinopathy—Detection of exudates in color fundus images of the human retina, IEEE Trans. Med. Imaging, Vol. 21, No. 10, pp. 1236–1244, 2002.

[19] Laÿ, B., Analyse automatique des images angiofluorographiques au cours de la rétinopathie diabétique, Ph.d. Thesis, Centre of Mathematical Morphology, Paris School of Mines, June 1983.

[20] Spencer, T., Phillips, R. P., Sharp, P. F., and Forrester, J. V., Automated detection and quantification of microaneurysms in fluorescein

angiograms, Graefe's Arch. Clin. Exp. Ophtalmol., Vol. 230, pp. 36–41, 1991.

[21] Mendonça, A. M., Campilho, A. J., and Nunes, J. M., Automatic segmentation of microaneurysms in retinal angiograms of diabetic patients, In: Proceedings of IEEE International Conference of Image Analysis Applications (ICIAP 99), 1999, pp. 728–733.

[22] Walter, T. and Klein, J. -C., Detection of microaneurysms in color fundus images of the human retina, In: Lecture Notes in Computer Science, Vol. 2526, Colosimo, A., Giuliani, A., and Sirabella, P., eds., Springer-Verlag, Berlin, pp. 210–220, 2002.

[23] Duda, R. O. and Hart, P. E., Pattern Recognition and Scene Analysis, Wiley-Interscience, New York, London, Sidney, Toronto, 1973.

[24] Ripley, B. D., Pattern Recognition and Neural Networks, Cambridge University Press, Cambridge, UK, 1996.

[25] Ward, N. P., Tomlinson, S., and Taylor, C., Image analysis of fundus photographs—The detection and measurement of exudates associated with diabetic retinopathy, Ophthalmology, Vol. 96, pp. 80–86, 1989.

[26] Phillips, R., Forrester, J., and Sharp, P., Automated detection and quantification of retinal exudates, Graefe's Arch. Clini. Exp. Ophthalmol., Vol. 231, pp. 90–94, 1993.

[27] Moreno Barriuso, E., Laser Ray Tracing in the Human Eye: Measurement and Correction of the Aberrations by Means of Phase Plates, Ph.d. Thesis, Institute of Optics, CSIC, Spain, June 2000.

[28] Osareh, A., Mirmehdi, M., Thomas, B., and Markham, R., Automatic recognition of exudative maculopathy using fuzzy c-means clustering and neural networks, In: Proceedings of Medical Image Underestanding and Analysis Conference, July 2001, pp. 49–52.

Chapter 8

Segmentation Issues in Carotid Artery Atherosclerotic Plaque Analysis with MRI

Dongxiang Xu,[1] *Niranjan Balu,*[2] *William S. Kerwin,*[1] *and Chun Yuan*[1]

8.1 Overview

Advanced atherosclerotic plaque can lead to complications such as vessel lumen stenosis, thrombosis, and embolization, which are the leading causes of death and major disability among adults in the United States. To reduce the healthcare costs, improved methods of diagnosis, treatment, and prevention of these kinds of diseases are very important [1].

Histological investigations have tied clinical complications to the existence of vulnerable plaques and have shown that certain plaques posed increased danger of causing clinical events. These vulnerable lesions are characterized by a large lipid core that is separated from the vessel lumen by a thin or weakened fibrous cap. Cap rupture is believed to lead to rapid plaque progression and/or patient symptoms [2, 3]. In recent years, many research work has been conducted in this area to find approaches that can effectively diagnosis and/or prevent the development of vulnerable atherosclerotic plaque. In diagnostic imaging, efforts have been made in at least two directions in the study of plaque features that are believed to be related to clinical outcome [4, 5]: the size of plaque and its tissue constituents. The focus of the first direction is more on the morphological features such as degree of vessel lumen narrowing and plaque's area/volume [6].

[1]Department of Radiology, BOX 357115
[2]Department of Bioengineering, University of Washington, Seattle, WA 98195

The second direction is trying to identify the tissue type distribution in plaque which is the only way to distinguish vulnerable plaques from stable plaques of similar size.

The motivation to study the constituents within carotid vessel wall is that evidence suggests different plaque tissue types yield different vulnerabilities to plaque rupture. Also, the location of plaque tissues, such as the distance to lumen, may play a role in plaque rupture. Thus, imaging and analysis techniques that are sensitive to plaque tissue types are needed and can subdivide a plaque into its constituent components. This chapter presents the postprocessing techniques developed for the identification of plaque constituents. In our laboratory, these techniques have been used to study the characteristics of the human carotid lesions that caused neurological symptoms [7] and of high cholesterolemia patients. Technically, magnetic resonance (MR) images obtained from advanced lesions in human carotid arteries present unique challenges:

1. *Small size of artery wall*: The carotid artery is usually less than 1 cm in diameter. Even if high-resolution imaging methods are used, practical limitations of MR scanners require the field-of-view of the image to be at least 13 by 13 cm, with a resolution of at most 512 by 512 pixels. Within these image dimensions, the subject, carotid artery, is normally about 40 by 40 to 100 by 100 pixels ranges. The comparatively small number of pixels makes the processing and analysis very challenging.

2. *Complexity of tissue constituents*: In our study, over 10 types of plaque tissues are identifiable within carotid artery wall, including lipid, hemorrhage, calcification, and fibrous tissue among the most clinically important. The plaque constituents may or may not be present, are generally unpredictable in terms of location, and can be intermixed.

3. *Difficulties in tissue separation*: Many studies have shown that any individual MR image can only distinguish between a limited numbers of plaque tissues, regardless of contrast weighting [8]. Therefore, a need exists to integrate the information obtained from several different contrast weightings, like T1W, T2W, PDW, and TOF so as to provide a single representation of all plaque constituents. To achieve this, multiple spectrum data segmentation is very critical in this study.

4. *Special processing requirement on fibrous cap*: The fibrous cap is a thin tissue layer that separates the lumen and other plaque tissues within the

blood vessel wall. Therefore, it is the critical feature in predicting the occurrence of rupture and monitoring the stability of patients' diseases. As a result, specialized segmentation techniques aimed specifically at characterizing the fibrous cap must be considered.

From image processing point of view, segmentation, the process of grouping image pixels into a collection of subregions or partitions that are statistically homogeneous with respect to one or more characteristics, such as intensity, color, texture, etc., has been a very important region analysis technique in medical image applications. The eventual goal of segmentation is to aggregate those neighboring pixels with similar features as a region and separate it from the others or the background in the image.

Since the partitioned regions sometimes do not contain any semantic meaning corresponding to the real physical object in image, image segmentation technique often serves as a low-level processing step in image-processing procedures. However, it is very crucial to the success of higher-level recognition process and plays as a deterministic role to the eventual performance.

There are three categories of MR data to be analyzed in this study: single contrast weighting gray level images, image sequences, and multiple contrast weighting images. Different from what are often analyzed in other applications, the images in this study are of lower quality due to the various noises involved in the imaging process. In addition, they are with more complicated structure than subjects are usually analyzed in other medical image studies, such as brain.

From the segmentation technique point of view, numerous approaches have been developed in the last decade, which are summarized in literature reviews [1, 9, 10]. These methods are implemented from different perspectives and have shown their successes by applying to various images. They can roughly be classified as following categories: region growing and splitting [11–14], edge detecting [15–18], random filed modeling [19–22], active contour modeling [23–27], as well as some hybrid of these methods [28–31].

After careful analysis of those existing approaches, however, it is easy to see that most of the algorithms presented in the literature can work well only for certain particular types of image and their performances are good if some image formation processes are taken into account in the segmentation procedure. Because of this intrinsic applicability limitation in the proposed models, it is

generally difficult to obtain ideal segmentation results when they are applied into other types of images. In recent years, some studies have been conducted to take advantages of a few algorithms to improve the segmentation performance, such as the wavelet MRF method [28], the region competition [29], and the Fuzzy Snake [32], etc. In this study, we develop our solutions following this problem-solving approach.

The research on segmentation techniques at an early stage was more on single frame gray level or monochrome images according to the survey provided by Haralick and Shapiro [10]. One category of these algorithms is region-based segmentation that includes region growing, splitting, and merging techniques. They generally use the intensity smoothness or similarity among neighboring pixels to find the regions. Another category of approaches find regions based on the discontinuities or edges in image. Since the closed contour is usually hard to be obtained in the edge map generated by various edge detectors, such as gradient operators or Canny edge operator [15], a tedious and more challenging linking procedure has to be employed to find closed region boundary. Active contour model (ACM)-based algorithms are a category of segmentation methods that search the contour of a particular object by minimizing a curve energy function. Bayesian inference based algorithms are another category of segmentation methods. They usually define the segmentation result image (a label matrix) as a sample of 2-dimensional random field and find the optimal solution by performing maximum *a posterior* probability (MAP) estimation. The label matrix is usually modeled as Markov random fields (MRF) [21, 22] and computed by means of clique potential functions according to its duality with Gibbs random fields (proved by the Hammersley–Clifford theorem [33]). In recent publications, most of the research works are focused on the accurate model description [20], optimized energy minimization searching algorithms design [19], as well as the performance improvement by introducing new models [28].

The study of image sequences segmentation can be regarded as an extended application area of single frame segmentation approaches. In addition to the segmentation results on each frame of the sequence, the correlation between adjacent frames is often considered, and hence, it makes the automatic/semiautomatic processing possible. Even though some approaches have been proposed before [34–37], the designs of algorithms are usually decided by the detailed correlated features in applications. In our study on the atherosclerotic

disease diagnosis, the 2-dimensional cross section images are obtained in a parallel sequence with small distance between each two along human's carotid artery.

While the study of gray-level image segmentation is still an active area of research, there is a growing need for solutions to partition multiple channel or multiple spectral images. In our study, multiple contrast weighting images on the same subject are used to identify tissue types. The other typical application areas of multiple channel image segmentation technique include remote sensing images, color images, and multiple modality medical images. Similar to the solutions for monochrome images, there also exist algorithms for multiple channel images segmentation based on edge detecting [38, 39], region growing [40, 41], and region splitting and merging [42]. However, they are facing the same problems as in monochrome domain. A new category of methods for multiple channel images is based on the histogram analysis [43] and clustering [44, 45] in multiple dimension data space. Because of the absence of spatial constraints in an image domain, the performances of such methods are often limited due to the existence of strong noises in images. The Bayesian-based approaches have also been proposed [46, 47]. They are basically extended 2-dimensional models to a multiple dimension data space. However, because of the dramatic increase of computation in the practical implementation they cannot be applied to those time-demanding applications, such as image database retrieval. In recent publications, Comaniciu and Meer [48] proposed a method that integrates the spatial constraints and feature domain cluster searching results to improve the segmentation results for color images.

The material in this chapter is organized as follows. Following this introduction, section 8.1 focuses on the segmentation technique for single contrast weighting (gray level) MR image. It covers Bayes's theorem, MAP estimation, MRF model, and the existing algorithms based on MRF. Finally, the QHCF algorithm is discussed and some experiments are conducted to analyze its performance. Section 8.3 covers the MR image sequence technique and an image segmentation framework, a MRF-based active contour model. It also includes a brief review of the traditional active contour model and its enhanced version, minimal path approach. Section 8.4 is about multiple contrast weighting MR image segmentation solutions. It consists of two parts. The first part describes a multidimensional MRF (mMRF) and its corresponding multidimensional QHCF based solution. The second part is about clustering-based segmentation method.

Section 8.5 introduces the specific segmentation methods that we use in fibrous cap analysis.

8.2 MRF-Based Gray-Level Image Segmentation

8.2.1 Introduction

In this section, we will discuss the segmentation techniques for gray-level image. This is because the subjects in our study, the MR images, are gray level intensity based, with pixel intensity within range 2^{12}–2^{16} defined by different MR scanner manufactures. In addition, the methods for gray-level image are usually the basis for processing of a MR image sequence and multiple contrast images.

Gray-level image segmentation techniques have been studied for years. Among the existing algorithms in literature, some are based on the pixel intensity distribution or histogram [49–52], some use region-based splitting/merging approaches [11–14], and some are derived from morphological operations [53, 54]. They have been successfully employed in many applications. However, the drawback of these algorithms is the poor performance in noisy environment. Some Bayesian inference based segmentation techniques [19–22, 55], using the MRF as image model to improve robust performance to noise, have been proposed in recent years and become very popular.

This section will focus on the MRF model and its application on gray-level image segmentation. An enhanced version of the Highest Confidence First algorithm is introduced.

8.2.2 Markov Random Field

MRF has become a significant statistical signal modeling technique in image processing and computer vision. Generally speaking, the MRF model assumes that the information contained in a particular location is affected by its neighboring local structure of a given image rather than the whole image. In other words, the estimation of pixel's properties, such as intensity, texture, color, etc., closely relates to a neighborhood of pixels, and this dependency can be characterized by means of a local conditional probability distribution. This hypothesis can

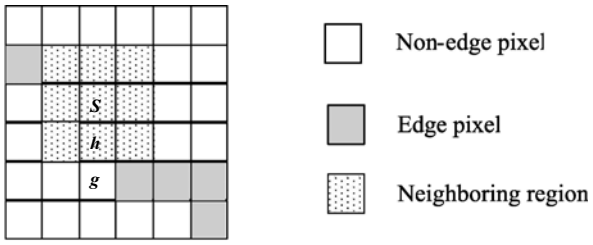


Figure 8.1: Illustration of MRF neighborhood and edge constraint. s and g are no-edge pixels belonging to different regions and h is an edge pixel within the s neighborhood.

reduce the complexity of the image modeling and provides a convenient and consistent way of describing the observed images.

8.2.2.1 Definition

Assume a two-dimensional random field is with size $\boldsymbol{I}$ by $\boldsymbol{J}$. For any pixel at location (i, j), $1 \leq i \leq \boldsymbol{I}$, $1 \leq j \leq \boldsymbol{J}$, its neighborhood, $N_{i,j}$, can be defined as

1. pixel $(i, j) \notin N_{i,j}$ and
2. for any pixel $(p, q) \in N_{i,j}$, there is $(i, j) \in N_{p,q}$

To illustrate the neighborhood, an example is shown in Fig. 8.1. The shaded region is a 3 by 3 neighborhood of pixel s. The size of neighborhood generally reflects how far a pixels surrounding region has an affect on it. This is a detail of the implementation of the algorithm and depends on the characteristics of the processed image.

Assume X is a two-dimensional random field and Ω denotes the set of all the possible samples of X. The definition of Markov random field is given as follow:

if X is an MRF, then for every $X_{i,j} \in X$ it must satisfy

$$\text{(i)} \quad P(X_{i,j}|X_{p,q} \quad \text{all } (p,q) \neq (i,j)) = P(X_{i,j}|X_{p,q} \quad \text{all } (p,q) \in N_{i,j}) \tag{8.1.a}$$

$$\text{(ii)} \quad P(X = x) > 0 \quad \text{for all } x \in O \tag{8.1.b}$$

Condition (i) is called the *Markovian* property that describes the statistical dependency of any pixel in the random field on its neighboring pixels. Under this constraint, only a small number of pixels within $X_{i,j}$'s neighborhood, $N_{i,j}$, instead of the whole image needs to be considered. Thereby, it reduces the

model's complexity significantly. Condition (ii), the *positivity* property, restricts all the samples of X to have a positive probability.

Although the Markov property is very useful for *a prior* model, the definition in Eq. (8.1) is not directly suitable to specify an MRF. Fortunately, the Hammersley–Clifford theorem [33] proves the equivalence between MRF and Gibbs random field (GRF), which states that a random field X is an MRF, if and only if, *a prior* probability $P(X)$ follows Gibbs distribution:

$$P(X = x) = \frac{1}{Z} \exp\left(-\frac{1}{T}U(x)\right), \tag{8.2}$$

where Z, *partition function*, is a normalizing constant. The parameter T is a constant used to control the peaking of the distribution. $U(x)$ is Gibbs energy function (or Gibbs potential). In the GRF model, a very important concept is called *clique*. The definition of *clique* C within the neighborhood of pixel s, N_s, is given as a set of points neighboring to each other, which satisfy

1. C consists of a single pixel, or
2. for $t \neq s, t \in C, s \in C \Rightarrow t \in N_s$.

The collection of all cliques is denoted by $C = C(s, N_s)$. An illustration of the types of cliques associated with C^1 and C^2 is shown in Fig. 8.2.

Based on this, the term of the energy function $U(x)$ can be expressed as

$$U(x) = \sum_{s} \sum_{C \in N_s} V_{sC}(x). \tag{8.3}$$

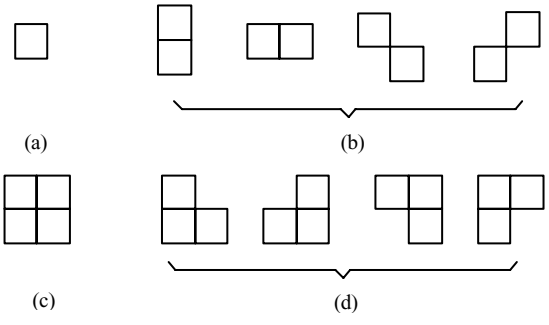


Figure 8.2: Illustration of all the possible cliques types associated with a 3 by 3 pixels neighborhood. (a) One pixel clique. (b) Two-pixel cliques. (c) Four-pixel clique. (d) Three-pixel cliques.

It is a summation of all the clique energy of each pixel along the whole image. $V_{sC}(x)$ is called clique energy function.

The assignment of clique energy is completely application dependent [33]. In our study, in order to obtain a precise model description, clique energy is calculated as a summation of two parts, pixel constraint and edge constraint [19]. The expression of clique energy function is written as

$$V_{sC}(x) = V_{sP}(x) + V_{sE}(x), \tag{8.4}$$

where the $V_{sP}(x)$ is the energy function derived by considering the spatial constraint of pixel s and its neighboring pixels. It is defined as

$$V_{sP}(x) = \sum_{h \in N_s} V_{sP}(s, h), \tag{8.5}$$

$$\text{with } V_{sP}(s, h) = \begin{cases} -\beta_1 & \text{if } x_s = x_h, \\ +\beta_1 & \text{if } x_s \neq x_h, \end{cases} \tag{8.6}$$

where x_s and x_h are the labels at location s and h. $V_{sE}(x)$ is the energy function with an edge constraint:

$$V_{sE}(x) = \sum_{h \in N_s} V_{sE}(s, h). \tag{8.7}$$

Assume h is an edge pixel within the neighborhood of pixel s (see Fig. 8.1),

$$V_{sE}(s, h) = \begin{cases} +\beta_2 & \text{if } x_s = x_g, h \in N_s, \text{ and } s,\ g \text{ are on different} \\ & \text{sides of an edge,} \\ -\beta_2 & \text{if } x_s \neq x_g, h \in N_s, \text{ and } s,\ g \text{ are on different} \\ & \text{sides of an edge.} \\ 0 & \text{otherwise,} \end{cases} \tag{8.8}$$

The introduction of edge constraint on energy expression is very straightforward. For two nonedge pixels, s and g, they are unlikely to be in the same region if an edge pixel, h, is in between of them (see Fig. 8.1). Compared to the energy function in traditional MRF models [20, 22], this additional edge constraint can provide more strict definition for energy function, so that the label-updating process can be more sensitive at boundaries of regions. Moreover, for those small regions with boundary points in the edge map (Canny edge detector is applied in this study), this edge constraint can also protect them from being merged with its large neighboring regions. In summary, the *a prior* probability for MRF can

be written as:

$$P(X = x) = \frac{1}{Z} \exp\left(-\sum_{s} \sum_{c \in N_s} \frac{V_{sP}(x) + V_{sE}(x)}{T} \right) \tag{8.9}$$

8.2.2.2 Maximum *a Posterior* Probability

Given an observed image Y, for any pixel at location s in Y, we assume it is with intensity y_s and with label x_s in segmentation label matrix. Then the *a posteriori* probability of segmentation result can be expressed as

$$P(X \mid Y) = \frac{P(Y \mid X)P(X)}{P(Y)}, \tag{8.10}$$

where $P(Y \mid X)$ is the conditional probability of the observed image given the scene segmentation. The goal of maximum *a posterior* probability (MAP) criterion is to find an optimal estimate of X, X_{opt}, given the observed image Y. Since $P(Y)$ is not a function of X, the maximization process only applies over the upper portion of Eq. (8.10), $P(Y \mid X)P(X)$. More accurately, given the observed image, the target of solving an MRF is to find the optimal state X_{opt} that maximizes the *a posterior* probability and take that state as the optimal image segmentation solution.

In this study, the conditional density is modeled as a Gaussian process, with mean μ_s and variance σ^2 for the region that s belongs to in the image domain. Thus, the intensity of each region can be regarded as a signal μ_s plus additive zero mean Gaussian noise with variance σ^2, and the conditional density can be expressed as

$$P(Y \mid X) \propto \exp\left(-\sum_{s} \frac{(y_s - \mu_s)^2}{2\sigma^2} \right). \tag{8.11}$$

By substituting Eqs. (8.9) and (8.11) into (8.10), the general form of the *a posterior* probability can be written as

$$\begin{aligned} P(X \mid Y) &\propto P(Y \mid X)P(X) \\ &\propto \left\{ \exp\left(-\sum_{s} \frac{(y_s - \mu_s)^2}{2\sigma^2} \right) \right\} \left\{ \frac{1}{Z} \exp\left(-\sum_{s} \sum_{c \in N_s} \frac{V_{sP}(x) + V_{sE}(x)}{T} \right) \right\} \\ &\propto \frac{1}{Z} \exp\left(-\sum_{s} \left[\frac{(y_s - \mu_s)^2}{2\sigma^2} + \sum_{c \in N_s} \frac{V_{sP}(x) + V_{sE}(x)}{T} \right] \right) \end{aligned} \tag{8.12}$$

Since Z is a constant, therefore, the a posterior probability can be simplified as

$$P(X \mid Y) \propto \exp\left(-\sum_s \left[\frac{(y_s - \mu_s)^2}{2\sigma^2} + \sum_{c \in N_s} \frac{V_{sP}(x) + V_{sE}(x)}{T}\right]\right) \tag{8.13}$$

Under MAP criterion, the optimal segmentation result should satisfy

$$\begin{aligned} X_{\text{opt}} &= \max_{X \in \Omega} \arg\{P(X \mid Y)\} \\ &= \max_{X \in \Omega} \arg\left\{\exp\left(-\sum_s \left[\frac{(y_s - \mu_s)^2}{2\sigma^2} + \frac{1}{T}\sum_{c \in N_s}(V_{sP}(x) + V_{sE}(x))\right]\right)\right\} \\ &= \min_{X \in \Omega} \arg\left\{\sum_s \left[\frac{(y_s - \mu_s)^2}{2\sigma^2} + \frac{1}{T}\sum_{c \in N_s}(V_{sP}(x) + V_{sE}(x))\right]\right\} \end{aligned} \tag{8.14}$$

We define the energy function of an MRF as

$$E(X) = \sum_s \left[\frac{(y_s - \mu_s)^2}{2\sigma^2} + \frac{1}{T}\sum_{c \in N_s}(V_{sP}(x) + V_{sE}(x))\right] \tag{8.15}$$

Therefore, the optimal segmentation searching is equivalent to the minimization process of the MRF:

$$X_{\text{opt}} = \min_{X \in O} \arg E(X) \tag{8.16}$$

8.2.2.3 Energy Minimization Method

The discussion in section 8.2.2.2 shows that in an MRF the optimal segmentation solution comes from finding the maximum of the *a posteriori* probability in Eq. (8.14), which is equivalent to the minimization problem of the energy function of Eq. (8.16). However, because of the large size of high-dimensional random variable X and the possible existence of local minima, it is fairly difficult to find the global optimal solution. Given the image size with $\boldsymbol{I}$ by $\boldsymbol{J}$ and the gray level for each pixel is $\boldsymbol{N}_{\text{g}}$, the total size of the random field solution space is $\boldsymbol{N}_{\text{g}}^{\boldsymbol{I}\times\boldsymbol{J}}$ that usually requires a huge amount of computation to find the optimal solution. For example, the size of MR image on carotid artery is usually 256 by 256, the gray level of each pixel is 2^{12}, then the number elements of the solution set is $(2^{12})^{256\times256} = 2^{786432}$, it is a prohibitive to be implemented in most interactive applications.

Some algorithms have been proposed to solve this problem in literature. Generally speaking, they can be classified into two categories. One category

is stochastic based, such as simulated annealing [56–58], Geman and Geman's Gibbs sampler [22], etc. They are all stochastic relaxation methods and theoretically can reach global minima by reducing the temperature slowly enough. Unfortunately, these algorithms normally require a huge amount of time and are intolerable for practical applications due to the extensive computation. The other category are deterministic methods, e.g., Besag's iterative conditional mode (ICM) [20] and highest confidence first (HCF) [21] proposed by Chou and Brown. These numerical approximation algorithms have much faster convergence speed and are often used as computationally efficient alternatives. ICM can be regarded as a special case of Gibbs sampler with an annealing temperature ($T = 0$). It may converge only to a local minimum and the results depend strongly on the initial estimate and order of updates, while the HCF algorithm, rather than updating the segmentation results via a raster scan order as that in ICM, selects the pixel with maximum confidence for updating. This overcomes the updating order problem in ICM and hence has the inherent advantage of faster convergence. In this study, we will only consider the algorithms on the second category because we promise to find practical solutions for segmentation requirements.

8.2.2.3.1 Simulated Annealing. Simulated annealing (SA), first introduced independently by Cerny [57] and Kirkpatric [58] as a stochastic solution for combinatorial optimization, simulates the physical annealing procedure in which a physical substance is melted and then slowly cooled in the process of constructing a lower energy configuration. The temperature control in the cooling process must be slow enough to allow all the substance to reach equilibrium state so as to avoid the defects. Metropolis *et al.* [59] proposed an iterative algorithm to simulate this annealing process. The searching of global minimum energy is controlled under a sequence of decreasing temperature T that satisfies the following rules:

(i) At higher T, a large increase in energy is accepted;

(ii) At lower T, a small increase may be accepted;

(iii) Near the freezing temperature, no increase is permitted.

Based on this, there are occasional energy ascents in the "cooling" process so as help the algorithm escape from local minima.

Suppose that the temperature parameter is at certain value T and the random field X starts from any arbitrary initial state, $X^{(0)}$. By applying a perturbation randomly, a new realization of the random field can be generated as $X^{(n+1)}$. The implementation of this perturbation varies in different optimization scheme. For example, in Gibber sampler, only one pixel is changed in each scan, while all the other pixels are kept unchanged. Generally speaking, the perturbation is usually very small so that $X^{(n+1)}$ is in the neighborhood of its previous state, $X^{(n)}$. The probability of accepting this perturbation is decided by two factors:

(i) Total energy change, ΔE, due to this perturbation.

(ii) Temperature T.

The definition of acceptance probability is defined as

$$P_{\text{accept}} = \begin{cases} e^{-\frac{\Delta E}{T}} & \text{if } \Delta E > 0 \\ 1 & \text{if } \Delta E \leq 0 \end{cases}. \tag{8.17}$$

It is obvious that the perturbations that lower the energy will be definitely accepted. However, when there is an increase of energy, the temperature parameter T controls the accepting probability in that given the same energy change ΔE, when T is with relative high value, the accepting probability is more than when T is relatively lower. Since this probability is based on the overall energy change, it has no dependency on the scanning sequence as long as all the pixels have been visited. In each iteration, this perturbing-accepting process will go on until the equilibrium is assumed as being approached (this is generally controlled by the maximum times of iteration). Then the temperature parameter T is reduced according to an annealing schedule and the algorithm will repeat the iterations for equilibrium searching as discussed above with the newly reduced temperature.

This annealing process will keep on going until the temperature is below the minimum temperature defined. Then the system is frozen and the state with the lowest energy is reached.

The annealing schedule is usually application-dependent since it is very crucial to the amount of computation in the stochastic relaxation process and the accuracy of the final result. Gemen and Geman proposed a temperature-reducing schedule that is expressed as a function of the iteration numbers:

$$T = \frac{\tau}{\ln(k+1)} \tag{8.18}$$

where k is the iteration cycle and τ is the constant. Even though this schedule can guarantee a global minimum solution, unfortunately, it is normally too slow for practical applications. Some other annealing methods [58] have been proposed to reduce the computation burden; unfortunately, it is no longer guaranteed to reach global minimum.

8.2.2.3.2 Iterated Conditional Modes. The goal of the ICM algorithm, proposed by Besag [33] in 1974, is to find a numerical approximation to further reduce the amount of computation produced by using stochastic relaxation. In recent publications, Pappas introduced an adaptive method [20] based on ICM and Chang *et al.* [46] extended it to color image segmentation.

In ICM algorithm, two assumptions are applied to the Markov random field:

(i) Given random field X, the observation components, Y_i, are modeled as the independent and identical distributed (i.i.d.) white Gaussian noise, with zero mean and variance s^2, and they are with the same known conditional density function $p(Y_s \mid X_s)$, dependent only on X_s. The conditional probability can be expressed as

$$P(Y \mid X) = \prod_{\text{all pixel s}} p(Y_s \mid X_s) \tag{8.19}$$

$$\text{where } p(Y_s \mid X_s) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y_s-\mu_s)^2}{2\sigma^2}} \tag{8.20}$$

(ii) The labeling of pixel s, X_s, depends only on the labels of its local neighborhood as

$$\begin{aligned} p(X_s \mid Y, X_r, \text{ all } r \neq s) &= p(X_s \mid Y, X_r, \text{ all } r \in N_s) \\ &\propto p(Y_s \mid X_s) P(X_s \mid X_r, r \in N_s) \end{aligned} \tag{8.21}$$

This is actually the Markovian property.

Under these assumptions, the conditional *a posterior* probability can be written as

$$p(X_s \mid Y, X_r, \text{ all } r \neq s) \propto e^{-\left(\frac{(Y_s-\mu_s)^2}{2\sigma^2} + \frac{1}{T}\sum_{c\in C_s} U_c(X)\right)}, \tag{8.22}$$

where N_s is the neighborhood of pixel s and C_s is the set containing all the cliques

within N_s. This equation shows that the local conditional probability depends only on X_s, Y_s, and N_s.

Based on these relations, the ICM iteratively decreases the energy by visiting and updating the pixels in a raster-scan order. For each pixel s, given the observed image Y and the current label of all the other pixels (actually only the neighborhood of pixel s), the label of X_s is replaced with one that can maximize the conditional probability as

$$X_s^{(n+1)} = \arg\max_{\text{all labels}} p(X_s^{(n)} \mid Y, X_r^{(n)}, \text{all}\, r \neq s). \tag{8.23}$$

Starting from the initial state, this algorithm will keep on running based on the procedure introduced above until either the predefined number of iterations is reached or when the labels of X do not change any more. Then it is regarded that a local minimum is reached.

Compared with the acceptance probability in SA method, only decrease of energy change is accepted in the ICM algorithm. This can be regarded as a spatial case when $T = 0$ because SA never accept any positive energy change when T is at zero temperature. This is why ICM is often referred as the "instant freezing" case of simulated annealing.

Even though ICM provides a much faster convergence than stochastic relaxation based methods, the solutions from ICM are likely to reach only local minima and there is no guarantee that a global minimum of energy function can be obtained. Also, the initial state and the pixel visiting sequence can also have effects on the searching result.

8.2.2.3.3 Highest Confidence First. *Highest confidence first* (HCF) algorithm, proposed by Chou and Brown [21], is another deterministic method for combinatorial minimization. Similar to the ICM algorithm, the two assumptions in Eqs. (8.19)–(8.21) also hold for HCF, and the minimal energy estimation, equivalent to maximizing the conditional *a posterior* probability of $p(X_s \mid Y, X_r, \text{all}\, r \neq s)$ for each pixel s, is also implemented iteratively.

The core feature of the HCF algorithm is the design of sites' label updating scheme. Assume L, with size N_L, is the set of the labels that are assigned to each site in the segmentation label matrix, which includes a special label $0 \in L$ to indicate the initial state of all sites. In the HCF algorithm, the site labeled with 0 is called "*uncommitted site*"; otherwise, it is called "*committed site*."

In the optimization process, once a label is assigned to an uncommitted site, the site becomes committed and cannot return to uncommitted state any more. However, a committed site can update its label through another assignment.

Instead of considering the energy change on an individual site through a raster-scan sequence as that in ICM, the HCF algorithm has the energy changes on all sites of the image within consideration via a measurement, called *confidence*, and updates the one with the highest confidence. The definition of confidence is derived from the local conditional probability in Eq. (8.22). Assume that the local energy function at site s is $E_s(x_s)$:

$$E_s(X_s) = \frac{(Y_s - \mu_s)^2}{2\sigma^2} + \frac{1}{T}\sum_{c \in C_s} U_c(X) \tag{8.24}$$

So the probability expression in Eq. (8.22) can be rewritten as

$$P(X_s = l \mid Y, X_r, \text{all r} \neq s) \propto e^{-E_s(x_s)} \tag{8.25}$$

where $l \in L$, is the segmentation label. The confidence measure $c_s(x)$ at site s is defined as

$$c_s(x) = \begin{cases} \max\{E_s(l) - E_s(x_i)\} & \text{when } s \text{ is committed with label } l, \\ & x_i \in \{x_n \mid n = 1, \ldots, N_L, x_n \neq 0, \text{ and } x_n \neq l\}; \\ \max\{E_s(0) - E_s(L_{\min})\} & \text{when } s \text{ is uncommitted,} \\ & L_{\min} = \arg\min_{x_i \in \phi}\{E_s(x_i)\}, \\ & \text{and } \phi = \{x_n \mid n = 1, \ldots, N_L, x_n \neq 0\}; \end{cases} \tag{8.26}$$

In Eq. (8.26), $L_{\min}$ is the one that can make the energy function at site s minimum among all the elements except 0 in label set L. When a site is uncommitted, it is obvious that $c_s(x)$ is always nonnegative. Label of the site with the maximum amount of confidence will be changed to $L_{\min}$. The exact purpose of this label updating process is to actually pick up the site where an appropriate label assignment can lower the energy of the whole image the most. In the meantime, the new label of this site influences the energy and confidence of its neighbors whose confidences need to be updated before the start of the next iteration. We can also regard the confidence as an indication that how stable the segmentation will be with the changing of the label at s. Obviously, the more stable is the label-updated image due to the change, the more confident we are that the label at s should be updated.

The HCF algorithm initially assigns all the sites with label 0 in the segmentation matrix. In each iteration, the algorithm will update only the site with the maximum confidence to the label that minimizes E_s the most. In the calculation of $c_s(x)$, only neighbors that are committed can affect the clique potentials. Therefore, once the label of a site changes, either from uncommitted state to committed state or from one committed label to another committed label, the confidence of its neighbors will have to be recalculated. The HCF algorithm stops when all the sites are committed and their confidence becomes negative.

Although it is claimed [46] that there is no initial estimate needed in HCF, some parameters, such as the mean value of each site, have to be provided in advance in order to get the confidence calculated when all the sites are uncommitted. For implementation, a heap structure is suggested in Chou's design that creates the highest confidence in the searching process faster.

Even though both the ICM and the HCF algorithms can only reach local minima of energy function, the results obtained with HCF are generally better than that with ICM [19, 46] because of its relatively optimal label-updating scheme. In HCF, the optimization result is independent of the scanning order of the sites. Even though this may lead to the increase of the computation amount and implementation complexity, the HCF algorithm is still regarded as a very practical solution with the fast growing of the processors' power.

8.2.3 QHCF Method

8.2.3.1 Algorithm

In the optimization procedure of the original HCF algorithm proposed by Chou and Brown [21], all the pixels are uncommitted and labeled 0 at the very beginning. A (prechosen) number of classes K are required in advance and the whole image is then presegmented (usually by K-means clustering) into K regions. The purpose is to estimate the mean of region that each pixel belongs to so as to initialize the energy of the whole image for HCF updating, and as the segmentation result is very sensitive to this initialization, the choice of K becomes very critical.

To overcome this problem, in the proposed QHCF algorithm, we first apply a Quad-Tree procedure to presegment the image instead of using K-means. The

advantages are as follows:

(i) There is no need to predefine the number of classes because the Quad-Tree algorithm can dynamically decide the partitions based on its splitting criterion.

(ii) K-means is a clustering method, in which the grouping is done in the measurement domain and has no spatial connectivity constraint to those pixels with the same label during the iteration process, while in the Quad-Tree splitting process, the grouping of pixels is done in the image spatial domain so that each region will not share labels with others.

The Quad-Tree procedure initially divides the whole image into four equal-size blocks. By checking the value of *region criterion* (RC) V_{rc}, each block will be evaluated whether it is qualified to be an individual region. The RC for each block B_i is defined as

$$V_{\text{rc}}(i) = \text{var}(B_i). \tag{8.27}$$

If V_{rc} is smaller than a predefined threshold T_{rc}, the block is regarded as a region. Otherwise, it will further be divided into smaller blocks by the same splitting method until every small block satisfies. The choice of T_{rc} is application-dependent and defines how uniform each region will be.

Compared with the other systematic initialization schemes, such as uniform grid initialization presented in [19], the Quad-Tree initialization is more accurate and efficient in adaptive detection of the initial regions. The reasons are listed as follows:

(i) The selection of initial sites is based on the consideration of pixels in the surrounding region due to the spatial constraint.

(ii) The points within the same region are not used for initialization repeatedly so that unnecessary computations can be avoided.

(iii) The iterative comparisons can be simplified during the HCF labeling process and the problem of "unlabeled small region" in [19] can be solved.

Moreover, by applying the Quad-Tree preestimation, the segmentation result can be more consistent than the uniform grid method when the initialization parameter changes in certain situations (this will be shown by experimental results at the end of this section).

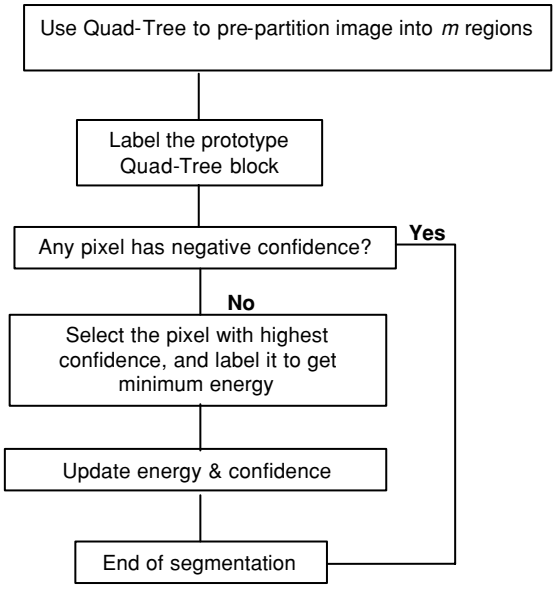


Figure 8.3: Procedures of the QHCF algorithm.

After the Quad-Tree initialization, in each region, the pixel with closest value to the mean of region intensity is selected as the representative and assigned a unique label; the others are all uncommitted and labeled 0. The algorithm will then start to update labels according to the procedures given in Fig. 8.3 until the energy of the whole image becomes stable. In this label updating process, the pixels are permitted to change only within the committed states or from uncommitted states to committed states. They are not allowed to become uncommitted from committed states.

To simplify the calculation, we assume variance with $\sigma^2 = \frac{1}{2}$. The local energy is normalized as

$$E_s(x) = (y_s - \mu_s)^2 + U_s(x), \tag{8.28}$$

and the confidence becomes:

$$c_s(x) = \begin{cases} \max\{E_s(x_k) - E_s(x_i)\} & \text{when } s \text{ is committed with label } k, \\ & x_i \in \{x_m \,|\, m = 1, \ldots, N_s, x_m \neq 0\}; \\ \max\{E_s(0) - E_s(L_{\min})\}, & \text{when s is uncommitted,} \\ & L_{\min} = \arg\min_{x_i \in \phi}\{E_s(x_i)\}, \\ & \text{and } \phi = \{x_n \,|\, n = 1, \ldots, N_L, x_n \neq 0\}; \end{cases} \tag{8.29}$$

representing the degree of confidence with which we can update the label of pixel s. Different from the definition in Eq. (8.26), for a committed site, the range for label searching is reduced to those existing within its neighborhood to decrease the confidence computation. For those uncommitted pixels at each Quad-Tree region, once any of their neighboring pixels is labeled, their energy and confidence will be affected. The confidence updating process needs to be conducted by applying Eq. (8.29).

After getting the new confidence of these sites, we search the whole image and select the one with the largest confidence as the next candidate to be updated. Any candidate site has one of three relations with its neighboring pixels: *isolated*, *surrounded*, or *partially surrounded*. When isolated (all neighboring pixels are uncommitted), the candidate pixel is given a new label different from existing labels. When surrounded (all neighboring pixels are committed), a unique label for this pixel becomes unlikely. We therefore select one label for this pixel from the labels owned by its neighbors according to

$$x_s - \min_{i \in L_N}\{E_s(x_i)\} \quad L_N \text{ is the set of neighboring labels of site } s, \tag{8.30}$$

making the energy of the candidate pixel become minimal. When the situation of partially surrounded (neighboring pixels are partially committed) occurs, we first consider it as a surrounded case and if the selected label x_k cannot satisfy $E_s(x_k) \leq \max\{E_s(0) - E_s(L_{\min})\}$, then a new label is assigned to this pixel. With this updating strategy, each region is entitled to have a unique label. This is different from the original HCF in which disjointed regions may share the same label. The advantage of the proposed QHCF is that the estimation of each region's mean value, μ_s, is decided solely by the sites in this region, resulting in more precise estimates during the label updating process.

As shown in Fig. 8.3, the segmented procedure stops when all the pixels possess negative confidence, where any change of a single pixel's label will increase the image's overall energy. However this does not guarantee that the image energy will converge to a global minimum. The tradeoff here is the processing speed because the adopted HCF algorithm can always finish in a finite time [21], providing a feasible solution to practical applications.

8.2.3.2 Experiments

Two experiments were designed to evaluate the performance of the QHCF algorithm. The first experiment used a phantom image to determine segmentation

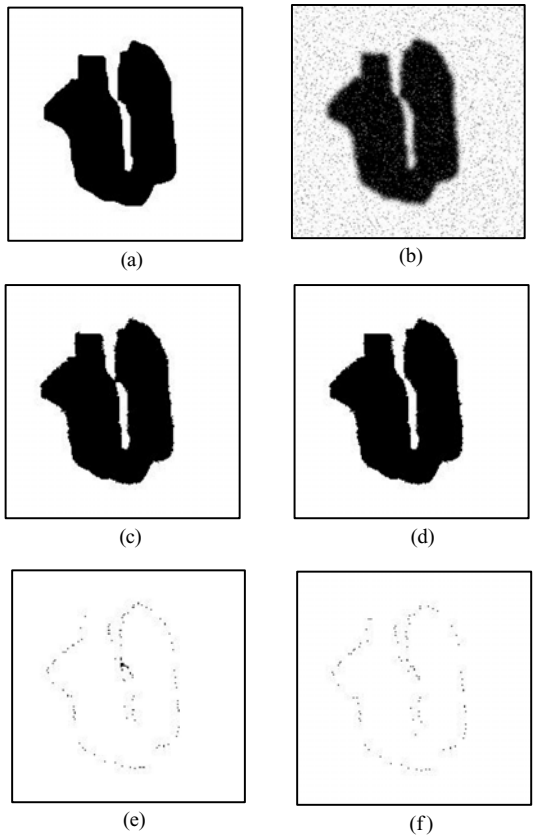


Figure 8.4: A phantom study of the MRF segmentation problem with adaptive ICM and QHCF. (a) Original phantom image. (b) Phantom is processed with additive Gaussian random noise. (c) Segmented image with Pappas' adaptive ICM algorithm. (d) Segmented result with proposed QHCF algorithm. (e) The difference between (a) and (c). Number of dark pixels is 120. (f) The difference image between (a) and (d). Number of dark pixels is 92.

accuracy. In this study, a phantom was used as the ground truth shown in Fig. 8.4(a). By applying additive Gaussian noise, a test image was created as shown in Fig. 8.4(b). It was segmented with adaptive ICM [20] and the QHCF algorithms individually. The results are shown in Figs. 8.4(c) and 8.4(d). Comparison of the segmented images with the original phantom images yielded difference images (Figs. 8.4(e) and 8.4(f)). QHCF had 92 pixel errors, while adaptive ICM had 120 pixel errors. Ten phantom image comparisons had been conducted and the QHFC algorithm sustained 24.7% fewer edge pixel errors than the ICM algorithm. We can also see from the shape of the ICM-segmented object that

merging of the two parts has occurred, while the proposed QHFC algorithm sustains the separation. This is due to the edge constraint in the QHCF energy function that makes it more sensitive at boundaries. However, in the analysis of the difference image (Fig. 8.4(f)) we note that most errors occur at the boundary of the segments creating a rough contour. Further work is indicated to solve this problem.

The second experiment was designed to evaluate the sensitivity of the segmentation result with differing initial conditions. We compared QHCF and uniform grid initialized HCF [19] (UGHCF) on human carotid MR images with a size of 128 pixels by 128 pixels. As UGHCF needs a predefined grid size, we chose 10, 20, and 30 pixels respectively. For the QHCF algorithm, the standard deviation of Quad-Tree region's intensity was used as RC V_{rc}, and the threshold was adjusted at 5, 10, and 20 intensity levels, respectively. Other constraint parameters, such as β_1 and β_2 have same values for the two algorithms. Figure 8.5 is an example of the segmentation result processed with the above initial conditions. Although the overall performance of the two segmentation results seem quite similar given various input RC values, QHCF gives more consistent results than UGHCF (for example, the partitioned regions within the white dotted line circles are stable in QHCF under differing initialization).

8.2.4 Discussion and Conclusions

In this section, even though experimental results demonstrate that the QHCF leads to better segmentation results over other approaches, including UGHCF and adaptive ICM, same as what occurred in other random field based solutions, the determination of parameters including T_{rc}, β_1, β_2, and T_{min} is also a difficult part of implementing the QHCF into real applications. This is mainly because the evaluation of segmentation result is usually application oriented, which highly depends on the verifications of object identification/recognition at the higher level in the system. Therefore, it is hard to find an ideal measurement in the lower level to feedback the segmentation performance. Even though there is no theoretical approach on the automatic optimal searching, some heuristic solutions can be adopted in our implementation.

(i) *Application-oriented empirical parameter selection*: For different applications, the requirements of segmentation results may be different

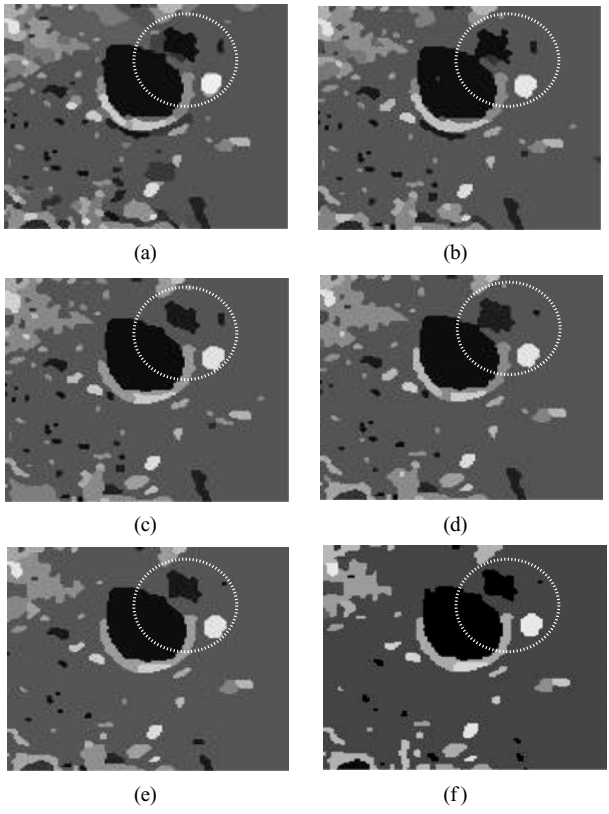


Figure 8.5: Comparison of initialization sensitivity between UGHCF and QHCF. The original is in Fig. 7(a). (a)–(c) Segmentation results under uniform grid initialization with grid size 10, 20, and 30 pixels, respectively. (d)–(f) Results *Reliability Criterion* threshold $T_{\text{rc}} = 5$, 10, and 20, respectively. The result within the white circle demonstrates the stability difference of the two algorithms with different initialization.

even with the same input image. Therefore, for a specific type of images, some empirical selections of parameters can be adopted. The parameters for two categories of images have been analyzed: one is about lumen segmentation with T1W MR images; the other is about the frame segmentation in videoconference clip. Table 8.1 shows the typical values of parameters in two applications for the QHCF algorithm.

Figure 8.6 is an example of the segmentation with different parameter combinations on T1W MR images. It shows that the parameter combination $T_{\text{rc}} = 10$, $T_{\text{min}} = 10$, $\beta_1 = 600$, and $\beta_2 = 100$ has better performance

Table 8.1: Parameters for QHCF algorithm

Empirical parameters	T1W MR Image	Video conference frame (QCIF)
T_{rc}	10	10
β_1	600	400
β_2	1000	600
T_{min}	10	10

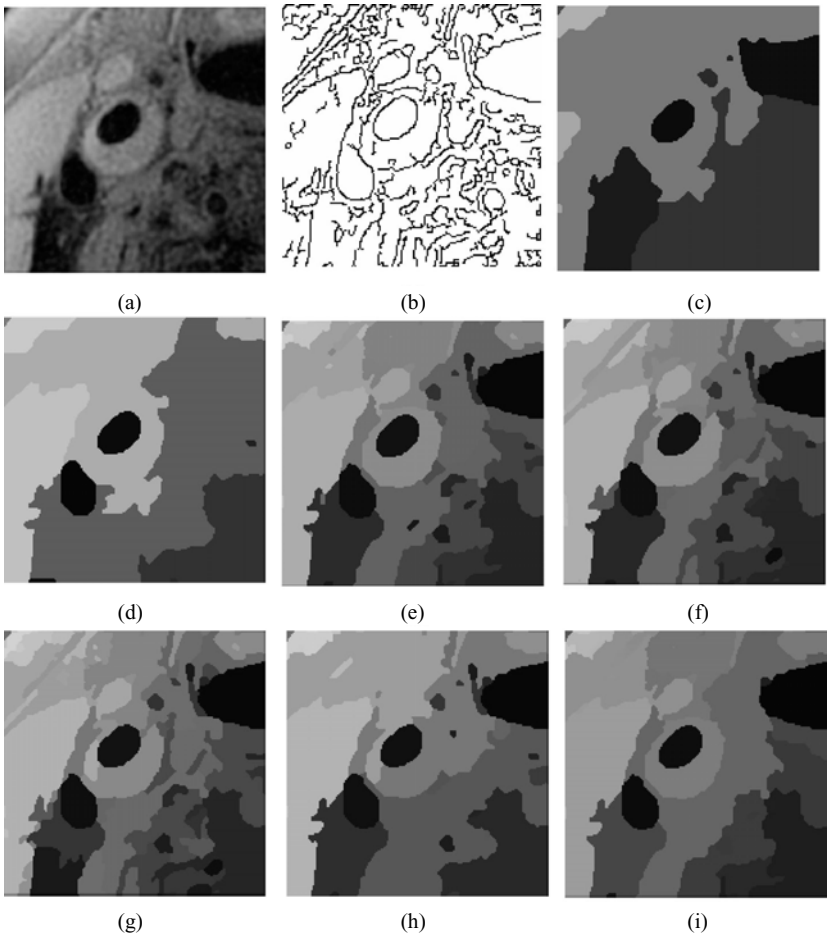


Figure 8.6: Segmentation of carotid artery lumen by QHCF, image size: 90 by 90, $T_{rc} = 10$, $T_{min} = 10$. (a) Original image. (b) Canny edge map. (c) $\beta_1 = 600$, $\beta_2 = 600$. (d) $\beta_1 = 600, \beta_2 = 800$. (e) $\beta_1 = 600, \beta_2 = 1000$. (f) $\beta_1 = 600, \beta_2 = 1200$. (g) $\beta_1 = 400, \beta_2 = 1000$. (h) $\beta_1 = 800, \beta_2 = 1000$. (g) $\beta_1 = 600, \beta_2 = 1000$, $T_{min} = 30$.

than others for lumen segmentation because all the typical regions, including lumens and blood vessel wall, are partitioned correctly. To further fine-tune the results, we increase the minimum region threshold as $T_{\min} = 30$ and an even "clear" result can be obtained as shown in Fig. 8.6(g).

(ii) *Supervised segmentation with interactive parameter selection*: Another solution for the parameter estimation in implementing QHCF algorithm into real application is to provide an interactive parameter selection interface for users. In our implementation, two modes were designed: auto and manual. In auto mode, a few empirical parameter combinations for typical applications are stored. The user can start with these predefined values and find one that can generate satisfactory results. Any further fine-tune of the segmentation performance can be achieved by switching to manual mode in which the parameters can be further adjusted.

8.3 MRF-Based Active Contour Model

8.3.1 Introduction

As discussed in section 8.1, most segmentation algorithms can perform well only on certain types of practical images because of the applicability limitation of each modeling or ease of use. In this section, we will introduce a flexible and powerful framework for general-purpose image segmentation. In the course of our investigation, we use the following assumption: *A successful segmentation is an optimal local contour detection based on an accurate global understanding of the whole image*. This assumption stems from the fact that the global information of an image is generally crucial in local object identification, automatic searching initialization, and energy optimization. Therefore, we focused our work in three parts:

(i) *Region segmentation of the whole image*: This provides a reliable basis for decision-making and subsequent processing.

(ii) *Local object boundary tracking*: This will optimally fine-tune the contour of the desired object region.

(iii) *Flexible identification mechanism*: This will bridge parts (i) and (ii) systematically and also be extendable to allow additional control functions or prior knowledge.

Our global-to-local processing logic is similar to the design concept of multiscale image segmentation algorithms [60]; however, subsequent processing is radically different. The multiple-scale based techniques use various processing methods applied to assorted resolutions of original images, normally from rough to fine. In the proposed framework, different models are employed to describe original image in different processing stages, from region based to pixel based and from global to local. Therefore, it is a hybrid solution by integrating the advantages of more than one algorithm.

In section 8.2 QHCF is employed as a deterministic implementation of MRF-based image segmentation. In this section, we use MRF as the basic model for the global region partition. To further track and enhance an object's boundary, we adopt the enhanced version of the active contour model (ACM) and minimum path approach (MPA) as the basic tracking method. A new scheme will be introduced to find ACM initial points based on the MRF region segmentation results so that it can automatically provide a reliable initialization. More specifically, unlike looking for those points in a potential field in previous solutions, we pick up the most reliable contour pixels from boundary points found by the QHCF and use the two-end-point based MPA to find the curves between every two adjacent ones. Then the whole boundary of object can be found by linking all these curve sections.

8.3.2 Survey of Active Contour Model

The active contour model, also known as *Snake*, was first introduced by Kass *et al.* [61] in 1988. In computer vision and image/video processing, it has been used as a very effective approach to implement contour tracking and shape feature extraction of interested object, and is also regarded as a successful deformable model in applications ranging from medical image analysis to video object manipulation. Basically, the development of active contour models has the following phases: classical Snake, geometric models, and minimal path approach models.

Roughly speaking, a Snake model can be regarded as a curve measured with an energy function. To track the contour of a desired object in an image, some

points or curves must be specified near the object's boundary initially. When the algorithm is applied, the Snake will "move" gradually toward the positions where the object's contour locates under certain constraints. This deformation process is generally conducted by iteratively searching for a local minimum of an energy function. However, a well-known problem of the classical Snake model is that it may be trapped into local minimal solutions caused by noise or poor initialization [62].

Another kind of active model is called the geometric active contour model that was first proposed by Caselles in 1993 [63]. It uses a geometric approach for the Snake modeling and applies the level set theory in the optimal curve searching. During the deformation process, the object contour evolutes and expands in the normal direction under certain constraints. Heuristic procedures are used to stop the evolution process when an edge is reached. The experiments presented in [63, 64] demonstrate better results than that was done with the classical Snake model [65, 66]. In 1995, Ceselles further improved the geometric model and transformed the object boundary detection problem as a path searching for minimal weighted length. This enhanced version is also known as the "geodesic model," which experimentally outperforms both the classical Snake model and the geometric model [67].

The minimum path approach, proposed by Cohen and Kimmel in 1996, is a state of the art solution in active contour modeling. It uses a path of cost as the interpretation of the Snake curve. The main feature of this method is that, given two prespecified end points, the global minimal path can be obtained. The energy optimization process is based on a numerical method proposed by Sethian [23] to find the "shortest path" in term of the global minimum of the energy among all paths joining the two end points. Compared with its previous versions of active contour modeling, MPA has the following advantages:

(i) It overcomes the local minimum problem in energy minimizing process.

(ii) It simplifies the initialization: only two initial end points are needed.

Nevertheless, this model still has some limitations in practical application. For example, the initial points must be precisely on the boundary of desired object to ensure the best contour searching performance. Therefore, human interaction is often required to accurately locate the initial points. Also, this model cannot address problems in which the shape of the desired object has topology change.

In the rest of this section, we will have a review of classical Snake model and the minimal path approach since they present the instinct spirit of this model and the state of the art of the optimal algorithm design.

8.3.2.1 Classical Snake Model

In the classical Snake model, a deformable curve is represented by an energy function whose local minima should be able to provide a set of alternative solutions based on the features of the object under investigation, such as shape, size, location, etc., as well as the surrounding image context. The optimization process is guided by energy minimization, which leads the deformable curve to evolve gradually from the initial contour toward the desired boundary of the closest object. The representation of the energy function contains two parts: internal energy E_{int} and external energy E_{ext}.

Generally, the internal energy E_{int} imposes only the smoothness constraint on the curve, such as elasticity and bending behavior, while the external energy E_{ext} is responsible for pulling the Snake curve toward the object's boundary. All these energies are formulated within an energy expression that is minimized by deforming the contour in an optimization process. The definitions are given as

$$E_{\text{Snake}} = \sum_{i=1}^{N} [E_{\text{int}}(i) + E_{\text{ext}}(\text{i})], \tag{8.31}$$

$$E_{\text{int}}(\text{i}) = \alpha_i \|v_i - v_{i-1}\|^2 + \beta_i \|v_{i-1} - 2v_i + v_{i+1}\|^2, \tag{8.32}$$

where N is the total number of Snake points; $v_i = (x_i, y_i)$ is a coordinate of the ith Snake point. Parameter α_i is a constant imposing the tension constraint between two adjacent Snake points. The stronger the α_i is, the shorter contour of the object will be obtained. Parameter β_i is a constraint tuning the bending among every three consecutive Snake points. Generally, the higher value of β_i will lead a smoother searched contour. In different applications where the object size and desired curvatures may vary, the values of α_i and β_i need to be adjusted. Obviously, it is unavoidable to have a procedure for parameters fine-tuning in practical applications. The definition of $E_{\text{ext}}(\text{i})$ is usually application- or feature-dependent. For example, $E_{\text{ext}}(\text{i})$ is often expressed as a form of image gradient function when boundary points of object are under searching.

The optimization process is implemented by iteratively finding a local minimum of the energy function given in Eq. (8.31).

Even thought the classical Snake model proposed has shown very good performance of object contour tracking in some applications, it has the following aspects that need to be further improved:

(i) *Optimization of the deformable curve*: Since the energy minimizing method in the original Snake formulation stops at a local minimum, it is obvious that numerical instabilities may be generated. Amini*et al.* [68] proposed to apply the dynamic programming approach to overcome this problem. Unfortunately, global minima cannot be guaranteed in his implementation [69]. In recent publications, Chandran and Potty [70] developed a method by imposing stronger conditions to force the optimizing process to avoid local minima. However, since these additional constraints have to be redesigned in each individual application, it is hard to be used as a general solution.

(ii) *Initialization of Snake model*: To track the contour of the desired object, some initial points or a closed curve are generally placed near the object's boundary in advance. This usually needs human's interactive mechanism, like *Snake pit* [71], involved to provide a reliable initialization. Otherwise, either a wrong target boundary may be tracked or the algorithm evolves with poor convergence [23]. To reduce this model's sensitivity and simplify its tedious initialization process, some methods have been proposed, such as "Snake growing" method by Berger *et al.* [72] and Fuzzy logic based framework by Eugene [73]. Cohen also introduced another method, called balloon force, to push the Snake curve outward from the center [74], which shows greater stability and faster convergence [75], by using the finite-element method. However, the limitations of this method is also very obvious, such as the location of initial points must be inside the desired object. Also, the optimal design of the balloon force is not closed.

In summary, even though some solutions have been proposed to resolve the drawbacks of classical Snake model, the main problems such, as initialization, topology changes, and energy minimization, have not been satisfactorily resolved.

8.3.2.2 Minimal Path Approach Model

Minimal path approach (MPA) is the recent version of the active contour model proposed by Cohen *et al.* [23] in 1997. Compared with its predecessors, the main improvements of this model are in the following two areas:

(i) Optimal curve searching with global energy minimization.

(ii) Simplified initialization.

Similar to other active contour models, the contour evolving process of MPA is also based on the minimization of an energy function. A potential field $\boldsymbol{P}$ is created based on the edge map of the original image, and the energy function is expressed as the integration over $\boldsymbol{P}$ along the curve. The goal of the optimization process is to find a curve whose energy function is minimum.

The definition of energy function is given as following: assume a pair of control points are $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, the energy of the curve $\boldsymbol{C}$ connecting this pair of points is

$$\begin{aligned} E(C) &= \int_{\Omega} \left\{ w \left\| \frac{\partial C}{\partial s}(s) \right\|^2 + P(c(s)) \right\} \mathrm{d}s \\ &= wL(C) + \int_{\Omega} P(C(s))\, \mathrm{d}s \end{aligned}, \tag{8.33}$$

where $C(0) = \boldsymbol{p_0}$ and $C(L) = \boldsymbol{p_1}$; w is a constant; s is the arc-length parameter; L is the length of the curve C; $\Omega = [0, L]$. The regularization term multiplied by the constant w measures the curve length and is used for smoothness control. By minimizing the energy function $E(C)$, the "shortest" path in the potential field will be picked up as the optimal solution among all the curves linking end points $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$. Similarly, we can find the curves between every two adjacent end points along the object boundary, so that the closed object contour can be obtained.

Nonetheless, the MPA method does not solve those problems appearing in the traditional active contour model completely. For example, MPA can find the optimal solution for curve searching between every two given points. This indeed can overcome the local minima problem. However, this also means a higher requirement to the location accuracy of initial points because the technique for searching these initial points is still open. Therefore, human interaction is often needed to locate the end points precisely.

Another drawback with MPA is that this approach lacks the topography handling ability. For some applications within our study, such as carotid artery lumen contour tracking in MRI sequences, the topology of blood vessel lumen in each cross-section images may change due to bifurcation, and it is impossible to apply MPA directly even though the initial points can be provided precisely. Therefore, a mechanism is needed to track the topology changes for automatic image processing.

8.3.3 MRF-Based Active Contour Model

The discussions in section 8.3.1 and 8.3.2 have shown that both MRF-based region segmentation and active contour model-based object contour tracking algorithms have unavoidable problems when applied in real applications such as automatic blood vessel lumen segmentation in MR sequence. In this section, we present a new framework for image segmentation, which integrates the advantages of MRF and active contour models and overcomes some of problems from each of them. Also, this framework is sufficiently flexible for additional prior knowledge to be plugged in so as to be adapted to various applications.

Note that the closed contour of the desired object is found by linking the sections of curves between each pair of adjacent initial points. These points are usually static in the optimization process and are very critical to the overall contour shape in the final result. To guarantee the accuracy of these initial points, human interaction is usually involved in MPA model based applications [73, 76, 77]. In the proposed framework, a new scheme is designed to automatically find these initial points, which will be called "control points" in the following context as a distinction of those defined in previous active contour models. The control points are identified from the boundary points of the desired object in the MRF based region segmentation result under *maximum reliability* criterion.

The basic structure of the framework is shown in Fig. 8.7. In step one, the QHCF procedure is applied to each of the input image to obtain a reliable region-based segmentation of the whole image. Then an object identification procedure is conducted. This is the step where prior knowledge of the desired object can be applied, and approaches such as the decision tree, fuzzy logic reasoning, neural networks, morphological operations, and so on can be employed in the design.

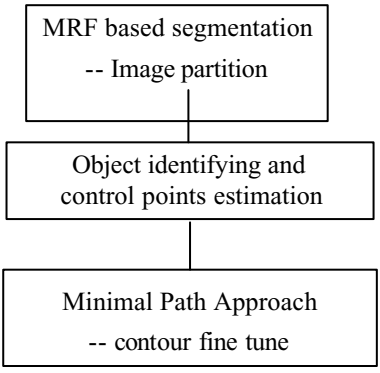


Figure 8.7: Structure of MRF-based Snake model.

Once the object of interest is identified, a procedure of optimal control point estimation operates among each section of its boundary points. Initialized by those control points, an instance of MPA model is setup and the optimal contour of the desired object is obtained.

8.3.3.1 Control Points Estimation

In the segmentation result obtained by the QHCF algorithm, the following information is available for further processing: region distribution, region intensity related properties (such as mean and standard deviation), and region boundaries. Although an MRF model can take into account the intensity continuity among neighboring pixels during the segmentation process, it imposes no constraint along the contour direction. Therefore, this problem of the QHCF method that there is no curve continuity constraint of object's contour during the optimization process makes the segmented object contour to be easily distorted due to noise. The experiment results in Fig. 8.5 have shown this drawback (rugged object boundary) that is unacceptable in some practical applications, such as quantitative medical image analysis and measurement.

In the proposed framework, a further fine-tune of region's boundary is accomplished by employing the MPA contour model [23]. To have an accurate initialization, it first needs to find the control points automatically.

As mentioned previously, the labeling process of each pixel in an MRF model is decided by the MAP, $\max\{p(x_i \mid y), i = 1, 2, \ldots, N$, where N is the number of labels$\}$. Based on this segmentation, the contour of an object can be easily found by searching region's boundary points. However, the experimental analysis of the

a posteriori probabilities of these contour pixels exhibit very large variations, indicating that the region boundary points found by the QHCF are not equally believed as the contour points. To reach the optimal contour tracking, it is necessary to select those most reliable ones as control points and search the other object boundary points by adhering to MPA constraints. The proposed solution for control point searching has the following steps:

Step 1. Locate the boundary points of the desired region based on the QHCF segmentation.

Step 2. Divide the region's boundary into sections.

Step 3. For each section, select the most reliable boundary point as a control point for MPA.

The selection process of step three is crucial to the success of the algorithm. Assume the boundary of the object of interest is divided into M sections and section m, $0 < m = M$, contains i_m total points. To simplify the problem formulation, we considered only the boundary points that have one adjacent region (they belong to another region). Suppose a particular boundary point is labeled p and its adjacent region's label is q, the *a posterior* probability of this point with label p and q can be expressed, respectively, as

$$\begin{aligned} p(x_s = p \mid y) &\propto \exp\left\{-(y_{i_m} - \mu_p)^2 - \sum_{s \in S_{i_m}} [U_N(x_s = p) + U_E(x_s = p)]\right\}, \\ p(x_s = q \mid y) &\propto \exp\left\{-(y_{i_m} - \mu_q)^2 - \sum_{s \in S_{i_m}} [U_N(x_s = q) + U_E(x_s = q)]\right\}. \end{aligned} \tag{8.34}$$

Assume this point belongs to region with label p, it is obvious that its *a posteriori* probability with label p should always have higher value than that with its adjacent region's label q. In a real image, like MR images or ultrasound image, noise affects the capturing process in boundary regions making the above assumption invalid. Distortion due to noise can blur edges and create a lack of separation in the *a posterior* probabilities of the true "edge points." To assure good measurement of the probability difference, we introduce the reliability of boundary points as

$$r(s) = 1 - [p(x_s = p \mid y) - p(x_s = q \mid y)]. \tag{8.35}$$

The value of the reliability is within the range [0, 1]. If s from the segmented

contour is more likely to be a boundary point, its *a posteriori* probability with label p and q will be quite similar. It then leads to the value of $r(s)$ being closer to 1 and makes point s more reliable. Therefore, in the control point searching process, we use *maximum reliability* as a criterion, which can be expressed as

$$s = \max_{0<i\leq i_m} \{r(s_i)\}. \tag{8.36}$$

The above criterion can be applied directly to the boundary points obtained with the QHCF algorithm because of the location and shape accuracy of the found object region. A further advantage of this accuracy is the solid foundation from which to do further work. This foundation is similar to the manual outline provided by the human operator for the traditional Snake algorithm. Consequently, use of the MRF-based segmentation result and the MAP criterion allows for an automatic initialization process that is relatively free of traps due to noise and spurious edges and has consistent reproducibility.

Step 2 addresses the selection problem of section number M and size of each section. Image quality and confidence of the contour points are determining factors in finding the solutions. For example, in our carotid lumen segmentation of MR images shown in Fig. 8.18, typical images generally needed 3–6 sections for contour tracking, while low-quality images required 8–10 splitting sections to track the whole blood vessel boundary. Object boundary corruption by noise results in more splits in the attempt to attain higher accuracy. The size of the object also is an important factor. Most of our studies contain objects sequestered within a square the size of 128 by 128 pixels. Division of the contour is accomplished by equal-length splitting. A more dynamic approach can be used in the case of a contour with noisier pixels resulting in more control points for ACM. The resulting curve will be noise-resistant and reliable. In addition, processing speed will be increased.

Step 1 is the most flexible and application-dependent of the three steps. It may also be totally eliminated in cases that target regions are already known. However, in most situations lack of advance knowledge of the exact location and spurious knowledge of the object's properties can be referenced as an additional constraint during the segmentation process. When this occurs an identification process can be designed based on the QHCF segmented regions to extract the boundary of the region-of-interest, which can then be used in further contour fine-tuning. An example is the lumen segmentation in a sequence of MR images.

The lumen may often be almost circular in shape and have a dark intensity. Applications in step one provide for the design of a decision tree to better identify the dark lumen from other regions in the QHFC segmentation.

In summary, the search for control points is the crucial step of this proposed framework. This step provides the bridge between the MRF segmentation algorithm and the active contour model. It decides the initialization accuracy, the key to the success of the minimal path approach. Finally, Step 1, being flexible, allows space for prior knowledge and integration of target constraints.

8.3.3.2 Contour Fine-Tune and Extraction

After finding all the control points in the M sections along the closed boundary, the outline of the desired object can be found similar to the human input initial points in the classical active contour models. However, compared with the human inputs, the identified control points are much more efficient and objective, especially in the situation where large amount of image sequences need to be processed.

Based on these control points, the complete contour can be found by applying the MPA algorithm to every two adjacent control points. To improve the performance in the optimization process, we dynamically frame the path searching range instead of applying it to the whole image. This can avoid the irrelevant sites in the image and hence reduce the computation. Assume a pair of control points are $\boldsymbol{P_0}$ and $\boldsymbol{P_1}$, in our implementation, the searching range is defined as a square containing $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, as illustrated in Fig. 8.8, in which $\boldsymbol{P_0}$ and $\boldsymbol{P_1}$ are the middle points of the edges. The shape of searching range is often decided by two factors:

(i) It must guarantee that the minimal path goes through this reduced search region.

(ii) The implementation of this region boundary control must be easy in case excessive computation is involved.

In practical applications, searching range is totally decided by the characteristics of the desired objects. A simple design is a band with certain width along the region boundary found by region segmentation. However, this may

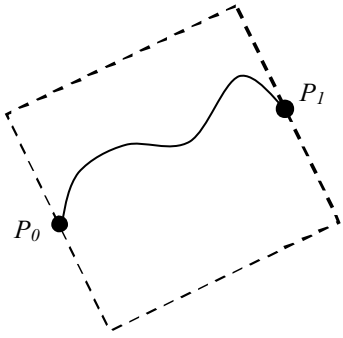


Figure 8.8: Illustration of the dynamic searching range frame.

involve a lot calculation to control its boundary. For simplicity and generality, we chose the square region to limit searching range. Another benefit of this restriction is that it can work as a control of the overall object shape and prevent the occurrence of "wild divergence" distorted by noise.

The procedure of region boundary splitting, control point searching, and curve fine-tuning is illustrated in Fig. 8.9.

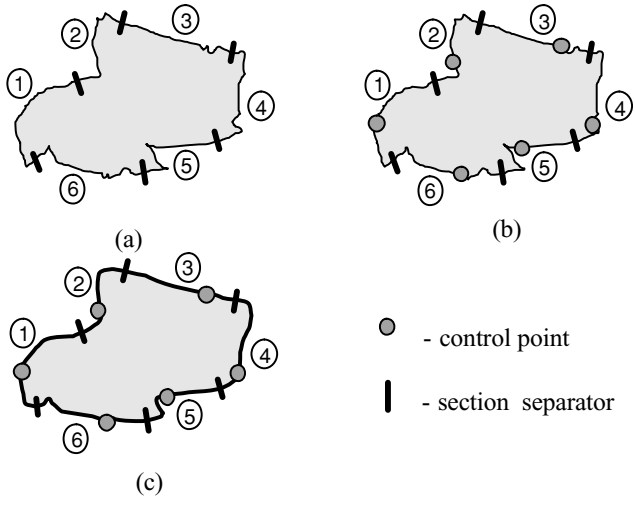


Figure 8.9: Illustration of the procedure used to apply the MRF-based active contour model on object boundary tracking. (a) The QHCF segmented region with the boundary divided into six sections. (b) In each section, a control point is searched based on *maximum reliability* criterion. (c) The final fine-tuned contour is found by linking the curves between each two adjacent control points, which are searched with minimal path approach.

8.3.4 Simulation Results

In this section, we present some experimental results of the proposed framework.

8.3.4.1 Segmentation of Single Image

First, a typical carotid MR image is shown in Fig. 8.10(a). Because of the noise or artifacts during imaging process, the intensity of lumen area is not uniform and there are some isolated bright spots inside. The QHCF algorithm was applied to this image and segmented it into many regions as shown in Fig. 8.10(c). From the result, we can see that the lumen segmentation is not affected by those bright spots inside the lumen area and most of the noise in the background have been suppressed. This is better than the result segmented with adaptive ICM algorithm shown in Fig. 8.10(b). By tracking the boundary of lumen region based on purely QHCF segmented results, we obtained the contour points of target region as shown in Fig. 8.10(d). It is obvious that some sharp corners on the top-left part of the contour and the bottom part are also not very smooth; this conflicts with normal observation of lumen shape in anatomy. In the next step, this contour was split into six equal sections and we searched the control points (see Fig. 8.10(e)) with maximum reliability criterion. The MPA algorithm was then used to track the whole contour and the result is shown in Fig. 8.10(f). Compared with the contour in Figs. 8.10(d) and 8.10(f), the effect of smoothness constraint in MPA algorithm is demonstrated. The two rough parts in Fig. 8.10(c) have also been fine-tuned.

8.3.4.2 Atherosclerotic Blood Vessel Tracking and Lumen Segmentation

Most existing active contour model based algorithms require the topology of the object to be known before the tracking action starts. Unfortunately, this requirement is difficult to be satisfied in some practical scenarios since the topography is often difficult to be predicted in advance. For example, in our study of carotid artery, the lumen bifurcates from one common carotid artery into internal and external carotid arteries at certain location along the image

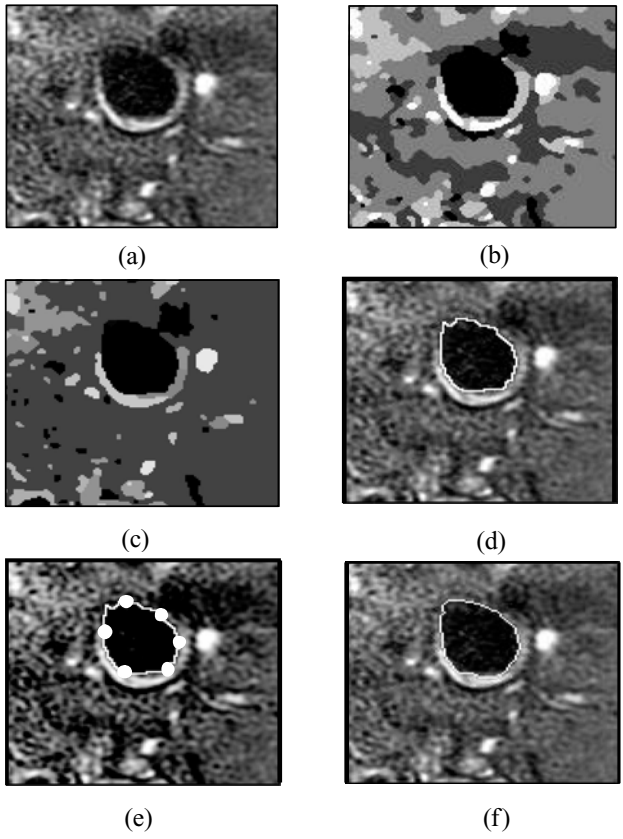


Figure 8.10: An example of lumen segmentation with MRF-based active contour framework. (a) The original MR image. (b) Segmentation results with adaptive ICM. (c) Segmentation results with the QHCF algorithm. (d) Rough lumen contour based on the QHCF segmentation result. (e) The six selected control points for MPA model initialization. (f) The fine-tuned lumen contour achieved by applying the MPA algorithm. In comparing (d) and (f), it becomes clear that contour tracking under the proposed framework results in superior smoothness control than with the MRF-based solution.

sequence. It is normally impossible to know where this kind of topological transformation happens.

8.3.4.2.1 Lumen Region Identification. To address the problem of topographical changes, we integrated prior knowledge into the control points

searching procedure. The bulk of our work is represented by the second block in the processing diagram shown in Fig. 8.7.

First, we model the MR image with MRF model and segment each of them into many regions by applying the QHCF algorithm. Since the number of lumen region may vary due to the bifurcation of carotid artery along the image sequence, a lumen identification process is indispensable before further plaque analysis. For each image, the lumen identification is achieved by letting all the segmented regions through a knowledge-based decision tree and picking up lumen region(s) of interest. The decision criteria are obtained by analyzing the statistical distribution of lumen region features based on prior knowledge in the test dataset. In the atherosclerotic blood vessel study, the following features are regarded critical for lumen identification:

(1) Region area C_{Area}

(2) Region average intensity $C_{\text{Intensity}} = \dfrac{1}{N}\sum_{n=1}^{N} I_n$ (8.37)

(3) Region circularity $C_{\text{Circular}} = \dfrac{4\pi C_{\text{Area}}}{L_{\text{Contour}}}$, (8.38)

where L_{Contour} is the length of region contour;

(4) Region location C_{Location}

The basic structure of the decision tree is shown in Fig. 8.11. For criteria C_{Area}, $C_{\text{Intensity}}$, and $C_{\text{Circularity}}$, statistical analysis of training MR image data required the use of two standard deviations as the satisfactory scale to make sure most of the variation range can be covered. For C_{Location}, it reflects the maximum radius of lumen center may locate in current slice away from the center of lumen in the previous slice. To reduce the computation in the identification process, the most distinctive feature of the target region is always analyzed first so as to decrease the number of candidates in the following criteria checking. In our study, the sequence is arranged as C_{Area}, $C_{\text{Intensity}}$, C_{Location}, and $C_{\text{Circularity}}$.

From above identification procedure, it can be seen that the accuracy of the low-level region segmentation plays a very important role in the topography detection. This can be achieved by using the QHCF algorithm. For the lumen

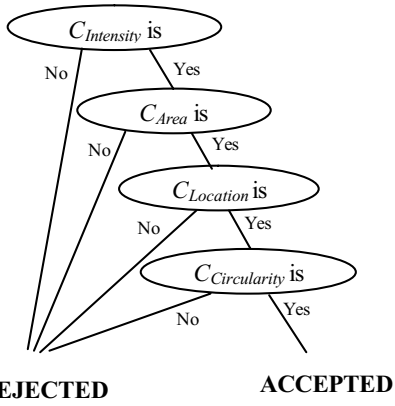


Figure 8.11: Diagram of the decision tree structure for lumen identification in MR image sequences.

identification step, however, if the choice of training dataset is sufficiently representative, those criteria will be fairly precise and hence make the decision result more stable. To further enhance the lumen tracking ability along image sequence, the location correlation of lumen regions between adjacent MR slices can be constrained by C_{Location}, which can further limit the searching range and therefore reduce computation.

8.3.4.2.2 Experiments and Discussion. In this study, 20 MR image sequences were tested with the proposed method. These images were scanned by a 1.5T SIGNA scanner (GE Medical Systems, 5.7 Echo Speed, custom made phased-array coils) in two imaging contrast weightings: T1-weighted (T1W) and 3D time-of-flight (3D TOF). Each mode produces 10 sequences with the bifurcation inside and 12 slices in each sequence.

It is well known that some feature characteristics appear altered by different modalities. The T1W sequence produces a lower intensity of the lumen signal. In 3D-TOF images, the intensity of the lumen is both higher and more uniform because of flow enhancement during imaging. For each contrast weighting, five sequences were selected for identification criteria estimation and the rest were used as testing data. Table 8.2 shows estimated criteria. The error rate is defined as the ratio between number of regions with error detection (misdetection or false alarm) and total number of lumen regions.

Table 8.2: Estimated criteria for decision tree

	T1W		3D TOF	
	Mean	SD	Mean	SD
C_{area}(pixel)	656[a], 364	105[a], 65	672[a], 371	97[a], 72
$C_{\text{intensity}}$	31	8	110	13
$C_{\text{circularity}}$	0.65	0.21	0.71	0.23
Error rate	3.3%		1.67%	

[a] For the common carotid artery.

For lumen contour tracking, prior use of the MPA model has provided more satisfactory results than MRF segmentation alone. Figure 8.12 is an example of the lumen segmentation procedure. A typical carotid artery MR image is shown in Fig. 8.12(a). Because of noise and artifacts during the imaging process, the intensity of the lumen area is not uniform and contains isolated bright spots. The QHCF algorithm was applied to this image with the result shown in Fig. 8.12(c) (note that lumen segmentation was not affected by bright spots in the lumen or background noise). The contour of interest region is shown in Fig. 8.12(d). It is obvious that some sharp corners on the left part of the contour and the bottom part are also not very smooth. Figure 8.12(e) shows the control points and the final MPA fine-tuned contour is demonstrated in Fig. 8.12(f). Figure 8.13 is an example of the blood vessel tracking in MR image sequence, with lumen bifurcation included.

Even though experimental results demonstrate good performance of the proposed framework, by analyzing those cases with error lumen identification it is found that the decision tree needs to be further enhanced so as to overcome the disturbance caused by random imaging artifacts in lumen region. Moreover, additional criteria and optimal decision strategies should also be considered in future research.

8.3.5 Conclusion

In this section, we discussed a framework, the *MRF-based active contour model*, for precise image segmentation and automatic contour tracking of image sequence data. It combines some of the most attractive features of random field

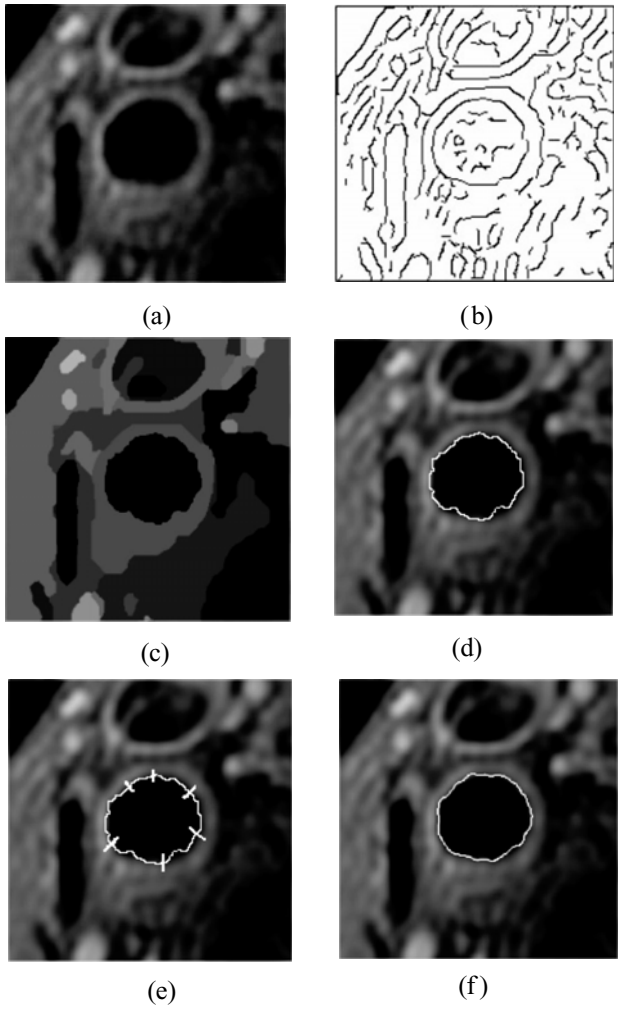


Figure 8.12: An example of the MRF-based active contour framework. (a) The original image of T1W MR image on carotid artery lumen. (b) Edge map by Canny edge detector. (c) Segmentation result of QHCF algorithm with $T_{\text{rc}} = 10$, $\beta_1 = 400$, $\beta_2 = 1000$, $T_{\text{min}} = 20$. (d) Lumen contour based on the QHCF algorithm. (e) Six selected control points. (f) Fine-tuned contour by applying MPA.

segmentation and ACM models. In addition, it is also very flexible and can easily include prior knowledge from various applications. An example of blood vessel tracking and lumen segmentation in magnetic resonance image sequences is studied and the experimental results have demonstrated very satisfactory performance.

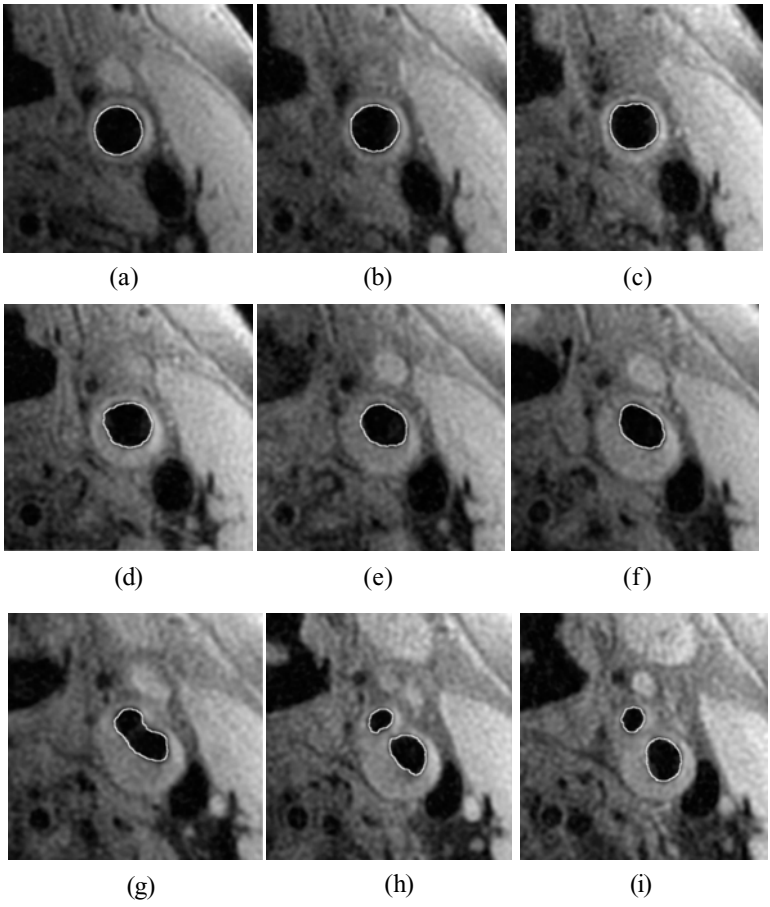


Figure 8.13: MR images of a human carotid artery from proximal common (a) through the bifurcation that occurs between images (g) and (h) to the distal internal and external carotids (j). Tracked lumen boundaries are visualized as distinct lines separating the lumen from adjacent tissues. This series also illustrates the topology change tracking ability of the proposed framework from the single lumen of the common carotid to the two lumens of the bifurcation, internal, and external carotid arteries. The location of real bifurcation happens between image (g) and (h). The closed bright curves along lumen boundary are the tracked contours. These results also demonstrate the topology change handling ability of the proposed framework.

8.4 Multiple Contrast Weighting MR Image Segmentation

8.4.1 Introduction

It is well known that an object viewed from multiple channels will generally convey more information than a single-channel observation [78–80]. A very successful application is remote sensing. Various sensors are designed to capture signals reflecting from surface of the earth in different bands. Since different objects on the earth have different spectrum profile, more details are usually detected by integrating the multibands information than that viewed with a single band. Similarly, in carotid plaque study, different imaging contrast weightings are often employed to detect the composition with blood vessel wall [81], and these multiple contrast weighting (MCW) techniques play a more and more important role in finding the different tissue types in the studied subject and generally can provide a more comprehensive view.

To achieve the goal of image segmentation and also to take advantage of the information with multichannel data, a multidimensional MRF (mMRF)-based solution will be first discussed in this section, which integrates the information from all different channels with a dynamical weighting. However, because of the intolerable amount of computation involved in the optimization process and intrinsic interspectral independency requirements in mMRF model, this technique becomes unsuitable to the requirements of interactive MR image analysis application. As a compromise, a robust cluster based segmentation algorithm is then put forward in our study, which is with faster segmentation speed.

8.4.2 Multidimensional MRF-Based Segmentation

In this section, we provide a multispectral image-segmentation solution based on mMRF model.

8.4.2.1 mMRF Model

Similar to MRF model discussed in section 8.2.1, our introduction of mMRF model is also based on MAP criterion. Assume the input images $I_1, I_2, \ldots, I_d$

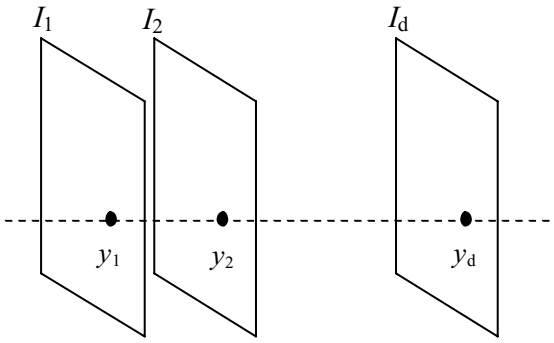


Figure 8.14: Illustration of d-dimensional image space.

are observed in d channels as illustrated in Fig. 8.14 and the label matrix of segmentation result is X. Then *a posterior* probability can be expressed in Eq. (8.39).

$$p(X \mid Y) = p(X \mid I_1, I_2, \ldots, I_d) \tag{8.39}$$

Another way to express the input data is to view them as d-dimensional vector $Y_s = [y_{s,1}, y_{s,2}, \ldots, y_{s,d}]^T$, where the value of ith dimension, $y_{s,i}$, represents the intensity at site s in image I_i. There is relation that

$$p(X \mid I_1, I_2, \ldots, I_d) = p(X \mid Y_1, Y_2, \ldots, Y_s), \tag{8.40}$$

where S is the total number of pixels in each image. Based on Bayes' theorem, *a posterior* probability is with form

$$p(X \mid Y) \propto p(Y_1, Y_2, \ldots, Y_s \mid X) p(X) \tag{8.41}$$

By assuming the conditional independence of each dimension given the segmentation result, the conditional probability can be expressed as

$$\begin{aligned} p(Y_1, Y_2, \ldots, Y_s \mid X) &\propto \prod_s p(y_{s,1}, y_{s,2}, \ldots, y_{s,d} \mid X) \\ &= \prod_s \left[p(y_{s,1} \mid X) p(y_{s,2} \mid X) \cdots p(y_{s,d} \mid X) \right] \\ &\propto \exp\left\{ -\sum_s \left[\sum_{i=1}^{d} \frac{1}{2s_{s,i}^2} (y_{s,i} - \mu_{s,i})^2 \right] \right\}. \end{aligned} \tag{8.42}$$

For prior probability, the neighborhood is defined as a block in mMRF model. Figure 8.15 is an illustration of 3 by 3 neighborhood in d-dimensional scenario. The black nodes are the pixels at location s in each channel and the gray nodes

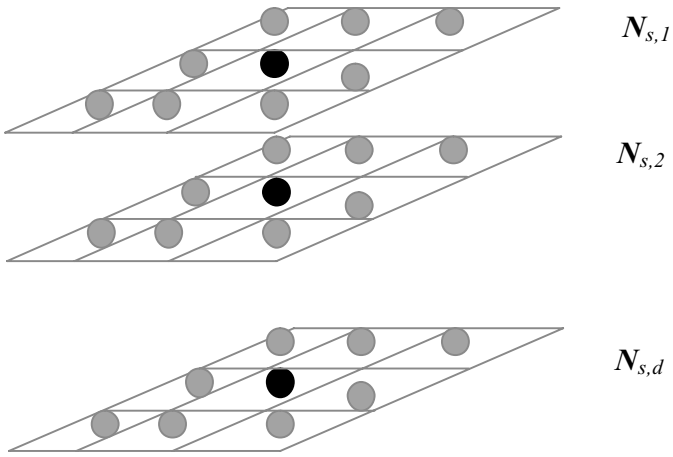


Figure 8.15: Illustration of a 3 by 3 by $\boldsymbol{d}$ neighborhood in mMRF model.

are the pixels in neighborhood. The prior probability of the whole image is

$$p(X) = \frac{1}{Z} \exp\left\{-\sum_s V_s(x)\right\} = \frac{1}{Z} \exp\left\{-\sum_s [V_{sN}(s) + V_{sE}(s)]\right\}. \tag{8.43}$$

Similar to the energy definition of monochrome image, the clique energy $V_s(x)$ for mMRF model at location s is a summation of spatial constraint energy $V_{sN}(x)$ and edge constraint energy $V_{sE}(s)$ in all channels. Their definitions are given in the following equations, respectively:

$$V_{sN}(s) = \sum_{i=1}^{d} V_{sN,i}(s), \tag{8.44}$$

$$V_{s\mathrm{E}}(s) = \sum_{i=1}^{d} V_{sE,i}(s), \tag{8.45}$$

where the $V_{sN,i}(s)$ and $V_{sE,i}(s)$ represent the component in the ith dimension. Compared to the energy function in traditional mMRF models [46, 82, 83], an additional edge constraint $V_{s\mathrm{E}}$ is added, which preserves the details of each dimension in the probability description and provides an even more accurate description of the energy function. This can makes the label updating process more sensitive at the regions boundaries.

Based on the above definition, we can find the *a posteriori* probability as

$$p(Y \mid X) \propto p(Y \mid X)p(X) \propto \exp\left\{-\sum_s \left[\sum_{i=1}^{d} \frac{1}{2s_{s,i}^2}(y_{s,i} - \mu_{s,i})^2\right] - \sum_s \left[\sum_{i=1}^{d} (V_{sN,i}(s) + V_{sE,i}(s))\right]\right\}, \tag{8.46}$$

and the energy function of the whole image is expressed as

$$E(X) = \sum_s \left[\sum_{i=1}^{d} \frac{1}{2s_{s,i}^2} (y_{s,i} - \mu_{s,i})^2 \right] + \sum_s \left[\sum_{i=1}^{d} (V_{sN,i}(s) + V_{sE,i}(s)) \right]. \quad (8.47)$$

Based on MAP, the optimal segmentation is

$$X_{\text{opt}} = \arg\max_X \{p(X|Y)\}, \quad (8.48)$$

which is also equivalent to the minimization of image energy $E(X)$.

$$X_{\text{opt}} = \arg\min_X \{E(X)\}. \quad (8.49)$$

8.4.2.2 The Algorithm

8.4.2.2.1 Definition. Even though many algorithms have been invented to minimize MRF's energy function as discussed in section 8.2.2.3, for mMRF model, only few algorithms have been proposed. The most recent solution published in literature was by Chang *et al.* [46], which is based on adaptive ICM method [84].

Since it has been shown in section 8.2 that the QHCF outperforms other algorithms, in this study, we extend it to multidimensional scenario. Its processing diagram is shown in Fig. 8.16. After having the images from all channels partitioned

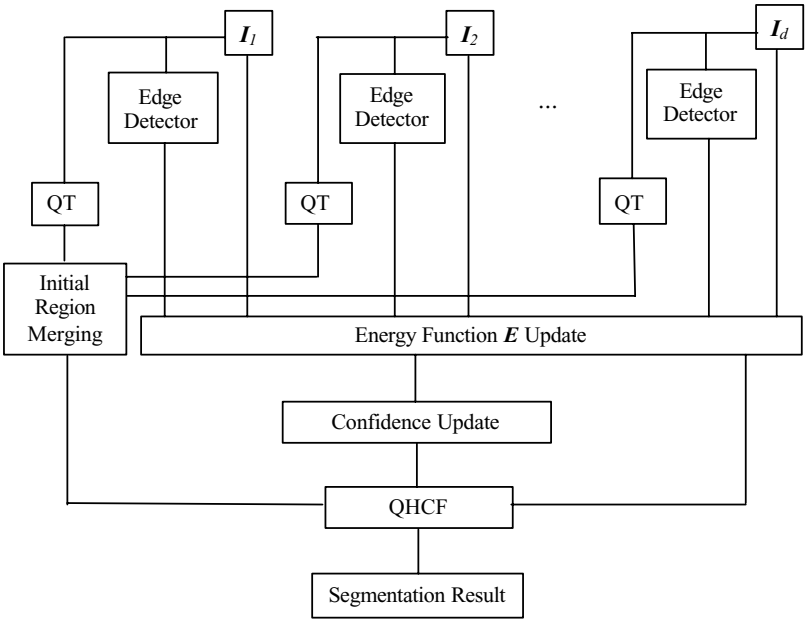


Figure 8.16: Design of segmentation algorithm for mMRF model.

with a Quad-Tree procedure, a "*initial region merging*" process is applied, in which a new region map is created as an initial segmentation result for the QHCF algorithm by combining all the regions found in different channels. The "merging" basically means the integration of region maps from all channels.

Under this presegmentation map, the definition of confidence for each pixel is the same as Eq. (8.29). However, the energy calculation is based on Eq. (8.47).

8.4.2.2.2 Dynamic Weighting. In multichannel data/image processing, different channels usually convey different amount of information. For example, in the soft tissue type identification with MR imaging [81], subjects are generally scanned with the multiple contrast weightings, such as T1-weighted (T1W), T2-weighted (T2W), proton density-weighted (PDW), and 3D time-of-flight (3D TOF). Since each contrast weighting imaging technique is sensitive only to certain tissue types, therefore, they usually contribute differently to the final decision when different tissue type is analyzed.

However, since there is no prior knowledge about the tissue type sensitivity layout in each channel, it is very impractical for human interaction involved in the segmentation process. In this study, a dynamic weighting system is proposed as a simulation of human's decision process, which automatically decides the weighting coefficient for the energy calculation among channels. In our implementation, two factors are significant in managing the dynamic weighting:

1. *Complexity factor* (CF): It measures the amount of details that each channel provides at a certain location. In the surrounding region of each location, we assume the complexity is proportional to the number of edges. The more edge points can be detected, the more details this channel can provide. Since Canny Edge detector [15] has been used successfully in the energy calculation, it is utilized in our implementation to generate edge map. Based on the requirements of segmentation performance, two ways are proposed to evaluate the complexity factor.

    (i) *Local CF*: The number of edge points within a local neighboring region in each channel.

(ii) *Global CF*: The number of edge points in the whole image in each channel (it is equivalent to local CF with neighboring range as the whole image).

Obviously, *global CF* is simple in terms of computation and represents the importance of each channel in a general sense. It is very efficient when one of channels plays a critical role in the segmentation process. On the other hand *local CF* is more complicated because it estimates the complexity of each channel at every location. However, it is very effective in preserving the segmentation details from each channel. Local CF is also very suitable to the situation where no prior knowledge of each channel's potential contribution to segmentation results.

2. *Weighting factor* (WF): It is used to calculate the exact weighting of each channel based on the measurement of complexity factor. Assume the complexity factor from each channel is represented as: CF_i, $i = 1, 2, \ldots, d$, the weighting factor is denoted as

$$WF_i = \frac{CF_i}{\sum_{i=1}^{d} CF_i}, \tag{8.50}$$

and the clique energy at each locataion $V_s(x)$ in Eq. (8.9) can be computed as

$$V_s(x) = \sum_{i=1}^{d} WF_i[V_{sN,i}(x) + V_{sE,i}(x)]. \tag{8.51}$$

8.4.2.3 Experiments and Discussion

Multiple contrast weighting MRI is an important imaging technique in clinical diagnosis. In this experiment, the mMRF version of QHCF algorithm (mQHCF) is applied on a set of *ex-vivo* atherosclerotic images.

Carotid endarterectomy specimens were scanned using a custom designed surface coil on a 1.5T GE SIGNA scanner with the following contrast weightings: T1W, T2W, PDW, and TOF. Based on observation and empirical knowledge, the tissues identified by T2W and PDW MR images are quite similar. To reduce computation complexity in the segmentation process, T2W images were removed, reducing the data space dimensions to three.

Figures 8.17(a)–8.17(c) show the original image scanned with T1W, T2W, and PDW, respectively. The circular shape object in the center of the image

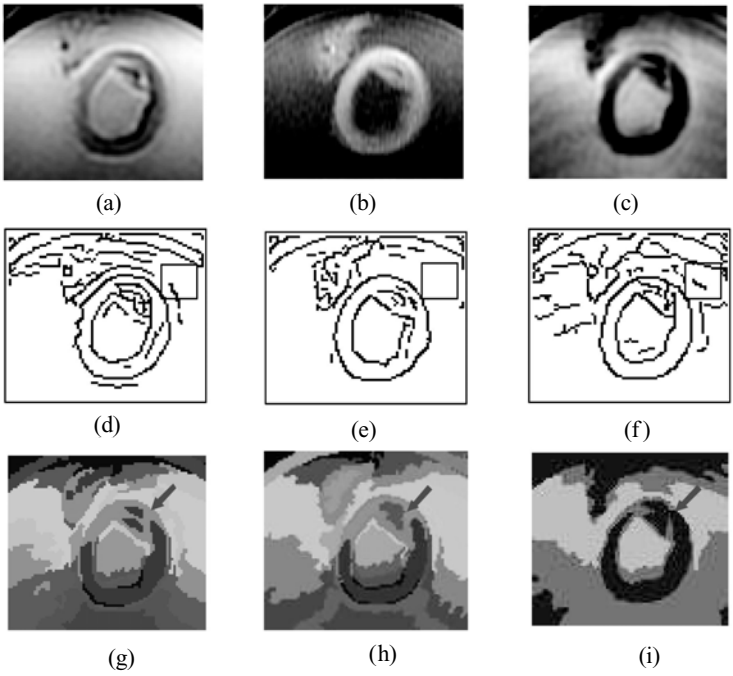


Figure 8.17: An example of *ex-vivo* atherosclerotic plaque segmentation with MCW MR images. (a)–(c) The region of interest in the original images in contrast weighing T1W, T2W, and PDW, respectively. (d)–(f) The corresponding edge maps by Canny edge detector. The square is an example of the region for local CF computing. (g) (color) Segmentation result by the proposed mMRF method with dynamic weighting. (h) (color) Segmentation result without dynamic weighting. (i) (color) The segmentations with QHCF on PDW MR images only. The red arrows point the regions showing the different revilements of details with MCW w/n dynamic weighting and SWC segmentation methods.

is an intersection of carotid artery. Figures 8.17(d)–8.17(f) are the Canny edge maps. In this experiment, we use *local complexity factor* to control the channel weighting; the small square represents the size of local range for complexity calculation in each channel. Figure 8.17(g) shows the segmented result. Compared with the segmented results shown in Fig. 8.17(h) (applying mQHCF only without dynamic weighting) and Fig. 8.17(i) (applying QHCF only on PDW channel) we can observe the fact that (i) segmentation result with MCW reveals more details than that with SCW; (ii) MCW with dynamic weighting can reveals more details (as shown in the area those red arrows pointing).

Table 8.3: Workstation configuration

Workstation	CPU	Speed	Memory	OS
Dell Precision 410	Intel PIII	600 MHz	256 MB	Windows NT 4.0

Another factor deciding the algorithm's performance is the processing time. Even though QHCF/mQHCF are deterministic implementations of the random field with finite optimization time, the computation is still a big cost for practical interactive applications. In mMRF, with the dimension expansion in the random field model, the amount of computation increases dramatically. This augment comes from two parts: (i) the computation used to calculate energy function and confidence for each location; (ii) the updating of its neighboring pixels. For other components in the optimization process, such as the updating of heap structure and searching of highest confidence, there is no big change involved. Since it is hard to compare the computation complexity theoretically, some experiments have been designed to compare the segmentation time for single modality image and multiple modality images. The experimental environment is set up as shown in Table 8.3.

The proposed algorithm was applied on 50 multiple contrast weighing MR images for each image size, the average time used for segmentation are given in Table 8.4. As a comparison, the average segmentation time for single modality image (T1W) is also listed. These experimental results indicate that the computation used for mMRF model is larger than that for single contrast weighting images.

Even though the discussion in section 8.4.2 has shown a solution for multiple channel image segmentation, there are some limitations in the applicability of the proposed mMRF based algorithms:

Table 8.4: Average segmentation time

Image size	Single contrast weighting image (sec)	Multiple contrast weighting image (sec)
128×128	13.202	92.104
256×256	37.531	244.328
512×512	69.459	517.163

(i) *Low processing speed*: From the experimental results in Table 8.4, the time used for multiple contrast weighting MR images is much longer than that for single contrast weighting ones, which is intolerable for practical interactive MR image analysis systems.

(ii) *Independency assumption*: In mMRF model, it is assumed that the signals in all the channels are regarded as independent as expressed in Eq. (8.42). However, in anthrosclerotic plaque study, even though the dependences among different contrast weighting images are unclear, there is no guarantee of their independency. Therefore, there is a risk that this assumption might be violated when new contrast weighting data is introduced.

8.4.3 Clustering-Based Method

Even though the mMRF may provide a robust solution to the segmentation problem of multiple spectral data, the intrinsic limitations of this model discussed in section 8.4.2.3 may hinder its further application. Therefore, there is a motivation to pursue a more relaxed and practical approach that can satisfy the following conditions:

(i) No presumption on the distribution function of the multiple dimension data.

(ii) Fast processing speed that is suitable for interactive applications.

In this section, a nonparametric clustering algorithm was developed to fulfill the above requirements and employed to process the multiple contrast weighting MRI images.

8.4.3.1 Introduction

Clustering is generally regarded as an effective technique for automatic data grouping based on a given similarity measurement. In the clustering process, discrete objects can be assigned to groups that have similar characteristics. The object can be single value, such as pixel's intensity in gray level image, or a vector in multidimensional data space. Generally, the clustering approaches usually consist of the following two steps:

(i) Cluster distribution or center searching.

(ii) Classification of dataset.

In the first step, the cluster center or the expression of density function is usually estimated so that the distribution of dataset can be clearly described. The second step is mainly on the dispatch of elements from the dataset based on certain criteria such as biggest similarity, shortest distance, maximum likelihood (ML), and K-nearest neighbor, etc.

8.4.3.1.1 Definition of Problem. First, we establish a clear definition of the studied problem. Suppose the number of MR imaging contrast weightings involved in the study is D, and images, $I_d, d = 1, \ldots, D$, are obtained from the same location of a patient's carotid artery. Assuming that there is no motion between these acquisitions and all the images have the same dimensions, $K = R \times C$, where R and C are the number of rows and columns of image. The data for each location $k, k = 1, \ldots, K$, can be expressed as $v_k = [v_{1k}, \ldots, v_{Dk}]^T$, $v_{dk} \in I_d, d = 1, \ldots, D$. v_{dk} is the intensity value at pixel location k in the contrast weighting image d. Based on this, a d-dimensional space with dataset $\boldsymbol{V} = \{v_k, k = 1, \ldots, K\}$ is created. Our goal is to construct a new segmentation image (same as the label matrix defined in HCF study), which contains uniform regions, some of which will be the plaque tissues of interest in the cross sectional image of the carotid artery.

8.4.3.1.2 Data Cluster Center Searching. Based on the above definition of MCW MR image segmentation problem, the two steps in the clustering approaches can be described more specifically as (i) to estimate the cluster centers of the data vectors in the d-dimensional space and (ii) to partition the K data vectors into clusters by mapping them to the "nearest" cluster center under certain rules so as to generate the composite segmentation image.

Cluster center searching in multidimensional space is also called the *multivariate location* problem, and numerous nonparametric methods have been proposed [85, 86]. Among them, the minimum volume ellipsoid (MVE) estimator by Rousseeuw [87] is one of the most well-known solutions. It is defined as the ellipsoid that satisfies the following two conditions: (i) covering at least h elements of dataset $\boldsymbol{V}$ in d-dimensional space; (ii) the minimum volume. Then the center of this ellipsoid is regarded as the multivariate location estimate, the

region with highest density. The MVE searching methods normally use randomly selected elements in $\boldsymbol{V}$ as the ellipsoids' initial centers. After inflating each of these ellipsoids until h elements are covered, the one with minimum volume is selected as a searched mode. Then we can remove the elements associated with this mode from dataset $\boldsymbol{V}$ and a new search is repeated until all the cluster centers are found. Although many approaches based on this multivariate locations estimator have shown its success in various applications [88], experimental results indicate that the performance of MVE decreases when the number of modes is greater than 10 [86]. The reason for this phenomenon is that the density definition in MVE presumes the multivariate normal distribution. Therefore, in the case of multiple modes, where no mixture of Gaussian distribution appears, MVE model will not be able to give an accurate description.

Another type of cluster searching techniques is called nonparametric estimation. The advantage is that they require no prior knowledge of the form of the density function in the search process. They can be applied to arbitrary distribution dataset. There are two main categories of methods for nonparametric density estimation, *Parzen window* and *k-nearest neighbors*. For the Parzen window approach, the kernel type needs to be given before it is applied. In the *k-nearest neighbor* method, the number of neighbors must be assigned in advance. Therefore, both require additional prior information. In addition, they are hard to optimally initialize.

To overcome these problems and provide more robust cluster estimation, a framework was proposed by Dorin [83]. It is based on the *mean-shift* algorithm, a nonparametric procedure to estimate density gradients. Although this method claims to avoid the drawbacks of most existing approaches, it still has the following weaknesses: (1) The initialization scheme cannot guarantee that all the cluster centers are under consideration in the search process because of its random tessellation selection and (2) because of the static size of the search sphere, the approaching speed may be slowed and the accuracy of cluster center estimation may be affected.

In this study, the data to be processed is MCW MR image, in which the distribution of data can vary arbitrarily between subjects. Since it is impossible to obtain the forms of the underlying density function, a nonparametric technique has to be employed for multivariate location estimation. To overcome the problems in Dorin's method, we first apply a preestimation of cluster distribution to guarantee that all the typical cluster centers are considered in the initial *center*

set. In addition, a dynamic sphere size is introduced to improve the searching performance.

8.4.3.1.3 Mean Shift Density Estimator. The mean shift algorithm was proposed by Fukunaga and Hosterler in 1975 as a "very intuitive" [89] estimator for data density. In 1995, Cheng generalized this algorithm and conducted a more rigorous study [90]. In this section, we will review its estimation of the density gradient in a uni-modal situation.

Assume $f(v)$ is the probability density function of a d-dimensional variable v. Suppose that centered with vector r, a sphere, S_v, with radius r is established. For any given vector within this sphere, y, the expected distance to sphere center v is

$$E[(y-v)|S_v] = \int_{S_V} (y-v) f(y \mid S_v)\, \mathrm{d}y = \int_{S_v} (y-v) \frac{f(y)}{f(y \in S_v)} \mathrm{d}y \quad (8.52)$$

With Taylor expansion, $f(v)$ can be expressed as

$$f(y) = f(v) + (y-v)^T \nabla f(v) \quad (8.53)$$

For $f(y \in S_v)$, when S_v is sufficiently small, it can be approximated as

$$f(y \in S_v) = f(v) V_{S_v}, \quad (8.54)$$

where $V_{S_v} = c \cdot r^D$ represents the sphere volume. Based on Eqs. (8.53) and (8.54), Eq. (8.52) becomes [83, 89]:

$$E[(y-v)|S_v] = \int_{Sv} \frac{(y-v)(y-v)^T}{V_{S_v}} \frac{\nabla f(v)}{f(v)} dy = \frac{r^2}{D+2} \frac{\nabla f(v)}{f(v)} \quad (8.55)$$

Expand the LHS of Eq. (8.55), the expected center of the sphere $E[v \mid v \in S_v]$ and v have the following relation:

$$E[v \mid v \in S_v] - v = \frac{r^2}{D+2} \frac{\nabla f(v)}{f(v)}. \quad (8.56)$$

For a given sphere, Eq. (8.56) shows that the difference vector between the local estimation of cluster center and sphere center is proportional to $\nabla f(v)$, the gradient of the density function at v. It is also reciprocal to $f(v)$. When it approaches the mode, $f(v)$ is generally large in value and $\nabla f(v)$ is small due to the slow increase, so a small mean-shift vector is applied. Compared to traditional density gradient searching techniques [86], in which only $\nabla f(v)$ is considered, the dynamically adjusted step size used in this method is more

accurate. Additionally, the mean-shift vector is always guaranteed to point to the direction of the maximum density. However, in the case when v is far away from the mode, the density change is often very small (close to uniform) within the small sphere under consideration. This may lead to the mean-shift algorithm failing to predict the correct direction and step size. Therefore, some measures need to be taken if the initial x is not at a location close to the mode.

8.4.3.2 Dynamic Mean-Shift Density Estimation

In the density estimation process, the mean-shift vector is very important to the search speed and accuracy of result. In previous methods [83], the size of the local estimation sphere is always fixed. This may not work well in the following two situations: (1) In a location which is far from the mode where the density distribution is relatively uniform, and if the sphere size is not large enough, the mean-shift vector may be misled and (2) when the search progress is close to the mode, the mean-shift vector needs to be sensitive enough to catch the local change. Therefore, if the sphere volume is too large, the local information cannot play the determinant role in the vector calculation and hence may affect the accuracy of center searching.

To solve these problems, we propose a *dynamic search range* with q levels: $r_1, \ldots, r_q$, where $r_i < r_{i+1}$. The values of r_1 and r_q come from prior analysis of MCW MR image data. In the search process, the initial radius starts from r_q(the largest radius). If the mean-shift vector is over the stopping threshold T_{stop}, it moves to the next location with the same sphere radius as the previous position. If the mean-shift vector is lower than T_{stop}, it uses the next smaller radius to calculate the mean-shift vector again. Once the mode is found, a small perturbation is applied [83] and the procedure is repeated to avoid a local maximum.

Given the initial center of a sphere at location v and starting with sphere radius index $i = q$, the proposed dynamic mean-shift algorithm is implemented in the following procedures:

1. Based on x and $r = r_i$, compute the mean-shift vector v_{ms}.
2. If $v_{ms} \geq T_{\text{stop}}$, $v = v + v_{ms}$ and repeat step 1; else if $i \neq 1$, $i = i - 1$ and repeat step 1 until convergence.

3. If perturbation has not been applied, add a small vector v_{pert} to the converged result and repeat steps 1 and 2, where $|v_{\text{pert}}| = |r_1|$ and its direction is randomly selected.

8.4.3.3 DMC-Based Segmentation

8.4.3.3.1 Cluster Center Searching. Based on the enhanced means shift density estimation, the steps for cluster searching are as follows:

1. Apply the Quad-Tree procedure to partition each contrast weighting image and assign a vector at the center of each region as a member of the initial *center set* $\boldsymbol{X} = \{\mathrm{x}_i, i = 1, \ldots, n\}$. In addition to the two constraints in previous method [83], the minimum neighboring points' distance and sphere density—a more strict constraint is applied on the pixel location in each contrast weighting image—the distance of the corresponding pixels' locations to any two elements x_i and x_j, cannot be too close.

2. For each point in the *center set*, apply the dynamic mean-shift search described in section 8.4.3.2 to find the candidate cluster centers in the d-dimensional data space.

3. Partition the *center set* elements into subsets in which the distances between points are within T_{sub}. Merging each subset by calculating the mean of elements in it, we arrive at a new center set $\mathbf{Y} = \{y_i, i = 1, \ldots, p\}$.

4. To validate the cluster centers, the constraint on the valley between every two elements in the *center set*, y_p and y_q, is applied [83]. Each point, for example y_r after a fixed interval along the line linking y_p and y_q is checked and the corresponding density $f(x_r)$ is estimated with Epanechnikov kernel [91]:

$$K_E(y) = \begin{cases} \frac{1}{2} c_d (d+2)(1 - y^T y) & \text{if } y^T y < 1 \\ 0 & \text{otherwise} \end{cases}. \tag{8.57}$$

Whenever

$$\frac{\min[f(y_P), f(y_q)]}{f(y_r)} \geq T_{\text{valley}}, \tag{8.58}$$

a valley is detected. If no valley is detected between y_p and y_q, the one with lower density will be removed.

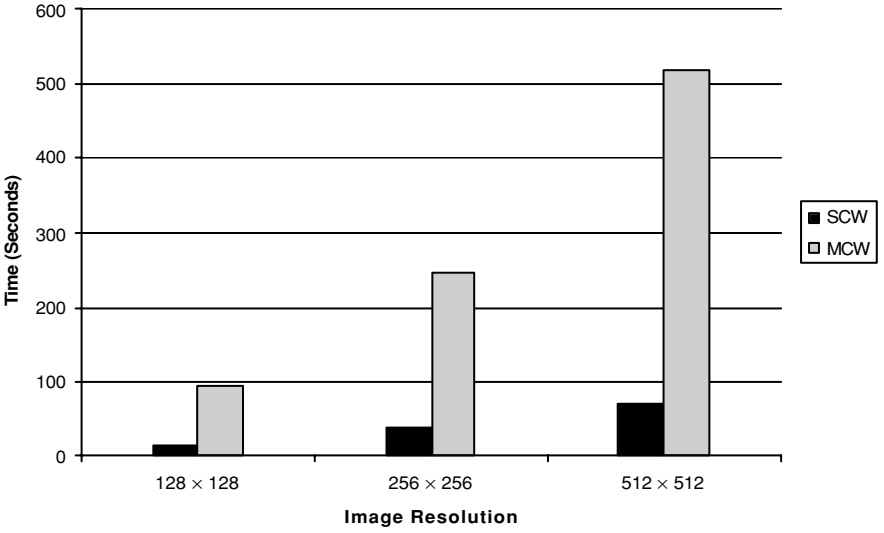


Figure 8.18: Average segmentation time comparison for multiple contrast weighting images and single contrast weighting images using MRF model based algorithm.

5. Using the elements in the *center set* as the cluster centers, the data in the d-dimensional space can be decomposed with the *k-nearest neighbor* approach.

8.4.3.3.2 Spatial Constrained Data Classification. The segmentation of MCW MR images is based on the cluster searching method described in section 8.4.3.3.1. After decomposing the d-dimensional vector space into p clusters, the segmented image can be derived by mapping each vector v in dataset $\boldsymbol{V}$ to a cluster center $y \in Y$. In addition, some image domain constraints can be further enforced so as to fine-tune the results. Figure 8.19 shows the segmentation procedure.

Note that the spatial constraint step is used to apply the spatial information in the image domain to the dataset dispatching process. Those isolated pixels that are unlikely to be independent regions are often merged with their neighbors. This decision is made based on the *a posterior* probability:

$$p(v_{ks} = L_i \mid v_{ks}) \propto \exp\left\{-(v_k - \mu_{L_i})^2 - \sum_{s \in S_{im}} [U_N(v_s = L_i) + U_E(v_s = L_i)]\right\}, \tag{8.59}$$

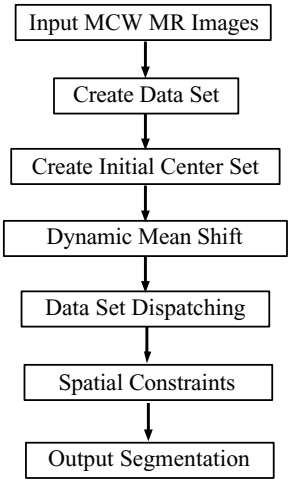


Figure 8.19: Processing flow chart of the proposed MCW MR image segmentation algorithm.

in which L_i and μ_{Li} are assumed to be the label and mean of neighboring region. U_N and U_E are the clique energy functions [92]. A MAP criterion is applied to find the most reliable neighboring region to merge into.

8.4.3.4 MCW MR Images Segmentation

Two categories of experiments are designed to apply the multispectral segmentation technique on MCW MR image. One compares the effectiveness of multiple contrast weightings in the MR image segmentation, and the other verifies the partition accuracy of typical plaque tissues.

To show the advantage of using multiple contrast weighting techniques over single contrast weighting, 20 cases of *ex vivo* MR images were obtained using T1W, PDW, and TOF. In each case, the described MCW segmentation approach was applied to the three contrast-weighted images with the result $\boldsymbol{I}_{\mathrm{MCW}}$. The single contrast weighting image in each case was also segmented using the same approach of assigning the dimension of vector space to one; the results are $\boldsymbol{I}_{\mathrm{SCW1}}$, $\boldsymbol{I}_{\mathrm{SCW2}}$, and $\boldsymbol{I}_{\mathrm{SCW3}}$. All other parameters remained the same. Preliminary work had indicated that composite segmentation might disclose information not available with single segmentation, this again proved to be the case. Figure 8.19 compares MCW to single contrast weighting segmentation results, whereby (a), (c), and (e) are the original images obtained with T1W, PD2, and TOF,

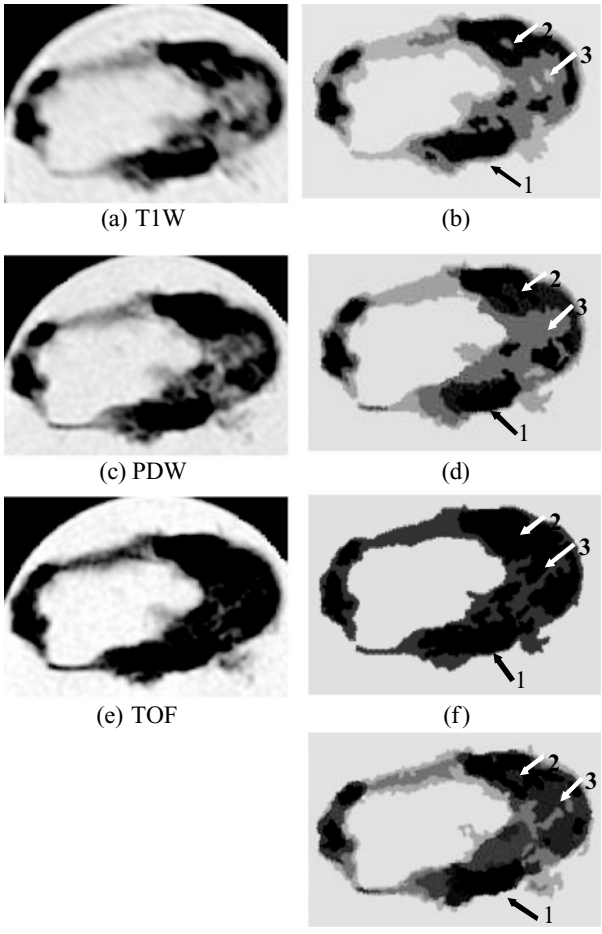
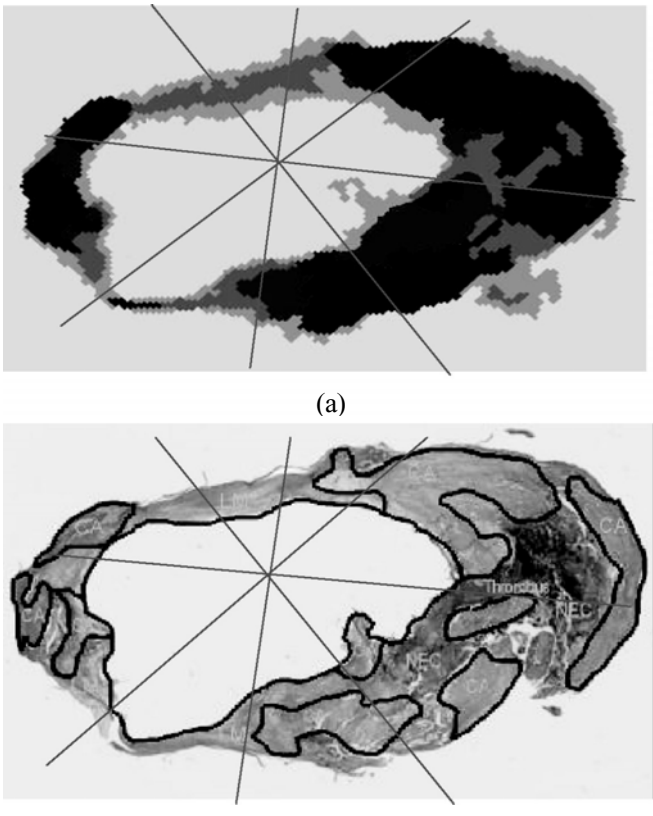


Figure 8.20: An example of MCW MR image segmentation. (a), (c), (e) The original contrast weighting images at the same location. (b), (d), (f) The corresponding single contrast weighing segmentations. (g) (color) Result with the proposed multiple contrast weighting segmentation algorithm.

respectively, and (b), (d), and (f) are the corresponding segmentation results using single contrast weighting only. Three distinctive differences are labeled in the images. Arrow 1 points to a region poorly segmented by the single T1W and TOF images. Arrow 2 shows a region that loses all detail in the TOF segmentation. Arrow 3 points to an area that by PDW segmentation shows no detail. However, the MCW-segmented image (Fig. 8.20(g)) retrieves and reveals these details by considering all contrast weightings in the segmentation process. Of 20 cases analyzed, 17 showed distinct differences occurring at more than two

locations between the MCW and SCW methods. Lower image quality (blurring, poor contrast, and imaging artifacts) was responsible for the poor segmentation result in the three remaining cases.

The second category of experiments was used to verify the partitioning of typical tissues of interest of the atherosclerotic plaque. Again we applied the above algorithm to *ex vivo* MR images of endarterectomy specimens. Using histology as the gold standard, we compared each segmented MR image to the best corresponding histology section. The carotid bifurcation was used as a landmark for location registration. On sections distant from the bifurcation, lumen size and shape and distinctive regions of calcification were used for matching. A coordinate system of eight segments was generated and applied to the matched histology and MR images. Figure 8.20 contains a pair of sample images.



(a)

(b)

Figure 8.21: An example of MCW MR images verification (**CA**: calcium, **LM**: loose matrix, **NEC**: necrotic debris). (a) (color) Segmentation of MCW MR images. (b) (color) Outlined corresponding histology section.

Table 8.5: Verification result

	Existence*		
Tissue type	Yes	No	Misdetect rate (%)
Calcium	69	2	2.8
Calcium (speckled)	18	2	10.0
Necrotic core	16	2	11.1
Necrotic core (mixed)	41	3	6.8
Foam cells	18	2	10.0
Fibrous plaque (dense)	165	8	4.6

Figure 8.21(a) is a MCW segmented MR image overlaid with the eight sector coordinate system. Figure 8.21(b) is the corresponding histology section stained with Mallory's Trichrome. Tissues of interest were outlined and labeled prior to matching.

A preliminary study of 22 matched MRI histology sections from eight patients were analyzed with results shown in Table 8.5. Typical tissues of the atherosclerotic plaque such as calcification, fibrous matrix, and mixed necrotic cores appear to have good agreement with histology. For those improperly matched cases, besides the inaccuracy of segmentation algorithm, the following may also be part of the reasons that affect the comparison results: (i) low-image quality (noise involved in the imaging process) and (ii) the deformation of plaques in the making process of histology section, including shrinkage. These are beyond the study of our research. However, further refinement of our technique may allow for better detection of the less discrete tissues such as loose matrix, speckled calcification, and intraplaque hemorrhage.

8.4.4 Conclusion

In this section, we investigated the segmentation algorithm based on multiple dimensional MRF model and clustering-based solutions, and introduced an effective approach for MCW MR image segmentation. This technique is based on mean-shift density estimation algorithm and was carefully designed to overcome the drawbacks in other existing methods. Experimental comparisons with histology section have demonstrated its successful performance.

For the processing speed of the proposed DMC-based approach, the same 50 multiple contrast weighing MR images with different image size were also

Table 8.6: Average segmentation time of DMC and mMRF

Image size	DMC (sec)	mMRF (sec)
128×128	8.608	92.104
256×256	19.140	244.328
512×512	27.937	517.163

used for testing. The comparison of the average segmentation time for DMC and mMRF approaches are shown in Table 8.6 which indicates that DMC uses much less time than mMRF.

In the validation experiments with histology sections, we can also note that poor image quality can reduce the accuracy of the proposed method by reading the cases that showed little agreement with histology. One hypothesis of this problem is that the poor separation of data in the vector space $\boldsymbol{V}$ makes the segmentation method unable to distinguish the different clusters.

8.5 Semiautomatic Detection of Fibrous Cap Status

Detection of fibrous cap status is crucial for understanding the disease status and prognosis of atherosclerosis. At the same time, fibrous cap segmentation is difficult because of resolution issues, registration issues, and the presence of artifacts. Hence a different approach is required to implement semiautomatic detection of fibrous cap status.

8.5.1 Importance of Fibrous Cap Detection

The development of a lipid core marks the development of an intimal xanthomata or fibrous streak into an atherosclerotic plaque. A thin layer of smooth muscle cells form a covering called the fibrous cap (FC) over the lipid core and separate it from the lumen [93]. Rupture of the FC in advanced lesions leads to thrombosis or intraplaque hemorrhage. Inflammatory destabilization of the

cap and subsequent thinning are prior events [94]. Thin FCs have been shown to be associated with symptomatic carotid vascular disease [95]. Studies of endarterectomy or postmortem histology identify such association retrospectively but methods of *in vivo* observation of FC status would enable prospective studies and lead to a better understanding of the pathogenesis. High-resolution MR imaging has shown promise in this regard. T2 [96], 3D TOF [7, 97, 98], and gadolinium-enhanced MR [99] have been used for FC imaging. Examination of multicontrast MRI with black blood (BB) sequences (T1, T2, PD) alongside 3D TOF has been shown to identify three different cap states: thick, thin, and ruptured [97]. A thick FC is considered to be stable while thin and ruptured caps are indicative of vulnerability. The presence of ruptured caps in MRI is highly associated with recent TIA [7]. MRI has shown a high sensitivity and specificity in identifying the three classes of FCs [98].

8.5.2 Identification of Fibrous Cap Status by MRI

Several image weightings (T1, T2, PD, 3D TOF) are used together by a radiologist to make the diagnosis of a FC [97]. A dark rim on 3D TOF is associated with a thick cap and a thin cap is associated with its absence. A ruptured cap is indicated by the absence of a dark rim in the presence of other markers like a focal contour abnormality or a bright gray region near the lumen [97] usually best seen in flow suppressed black blood images. Figures 8.22–8.24 show typical appearances of thick, thin, and ruptured caps, respectively. It has to be noted that various cap states can occur within the same slice.

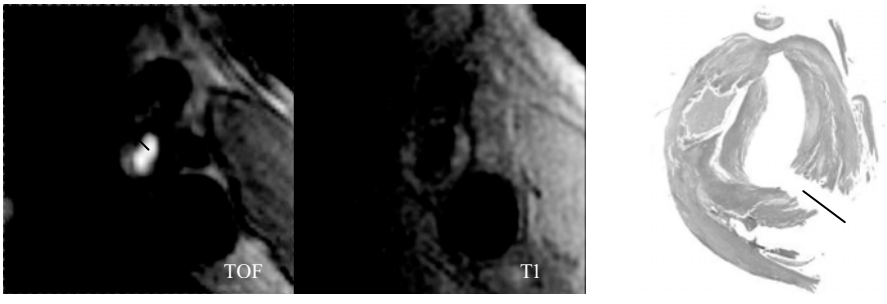


Figure 8.22: Thick cap—presence of dark rim on 3D TOF due to a thick cap (arrow). The site of apparent rupture on histology (color) is artifactual and caused by surgical incision (arrow).

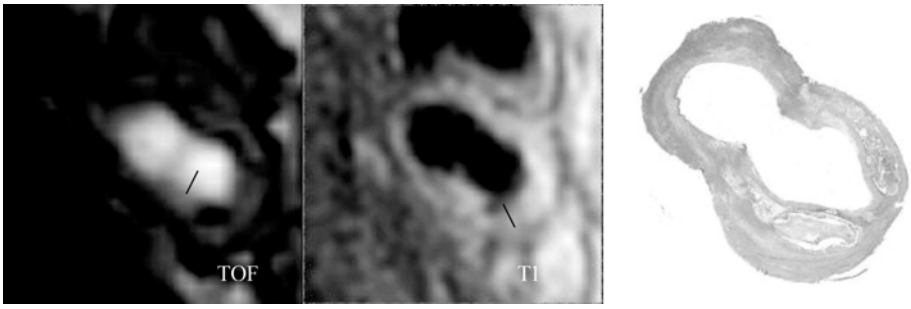


Figure 8.23: Thin cap—absence of dark rim on 3D TOF due to a thin cap (arrow). The apparent focal contour abnormality on T1 (arrow) is due to calcium and not a real contour abnormality.

8.5.3 Challenges in Identification of the Fibrous Cap

Fibrous cap thickness is in the order of a few tenths of a mm while the maximum resolution by MRI by the current protocol used is 250 μm $\times$ 250 μm (interpolated by zero filling) while the native MR resolution is around 500 μm. In spite of this resolution, 3D TOF seems to be able to detect the cap status [7, 97, 98]. When using multicontrast MR images for segmentation, registration accuracy becomes very important. Since the characteristics for FC detection by MRI occupy a few pixels around the bifurcation registration for multicontrast segmentation is difficult since different weightings will overestimate or underestimate the lumen size by some pixels. The algorithm outlined below takes into account the above factors.

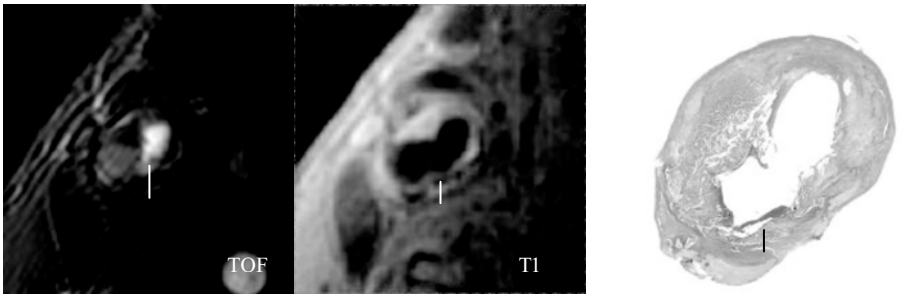


Figure 8.24: Ruptured cap—absence of dark rim on 3D TOF and focal contour abnormality on T1 due to a ruptured cap (arrows). The site of rupture on histology (color) is indicated by an arrow.

8.5.4 Semiautomatic Detection

It has been our experience that the following characteristics are primarily used to differentiate stable from unstable caps using multicontrast MRI:

1. Appearance of a dark rim in 3D TOF images implying the presence of a thick cap.
2. Focal contour abnormality best observed in black blood images and implying a rupture or erosion of the fibrous cap.

The above two characteristics are used by the algorithm since they are the primary distinguishing characteristics. Absence of a dark rim is taken to indicate a thin or ruptured cap. Ruptured caps can be differentiated from thin caps by the presence of a focal contour abnormality. Other factors such as the presence of calcium near the lumen surface [98], flow abnormalities [7, 98], and intraplaque hemoharrage [97] may affect the correspondence between FC status and the dark rim but are not currently taken into account by the algorithm. To perform the FC evaluation, matched 3D TOF images and one black blood weighting are used by the algorithm to identify plaque status (Figure 8.25). Parameters for the dark rim are measured from the TOF image and those for focal contour

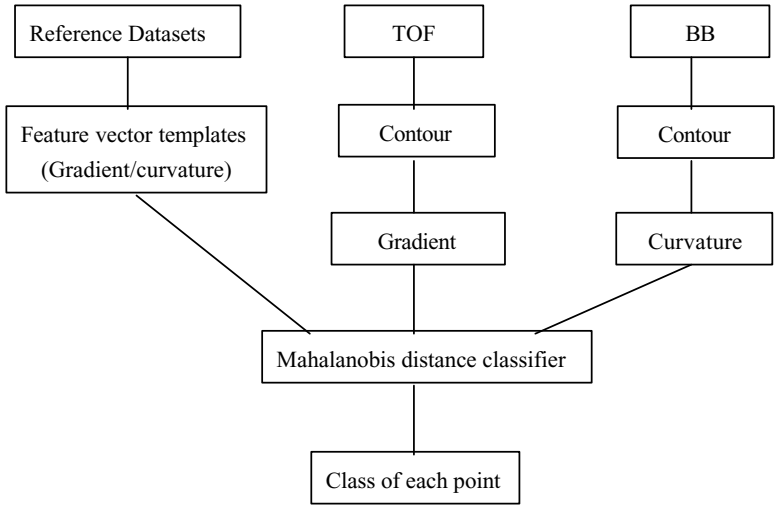


Figure 8.25: Schematic of algorithm—feature vector templates derived from reference datasets are used to classify points on the lumen contour based on their mahalanobis distance in the feature vector space.

abnormality are measured from one BB weighting. An operator draws lumen contours on both TOF and BB images by a Snake algorithm with appropriate weightings for the image energy term [100]. This step is the only semiautomatic step and the remaining steps do not require manual intervention.

In this approach, the algorithm classifies all points around the lumen contour, although the FC technically covers only the lipid core. With this approach, the human operator does not need to identify the body of the plaque and hence allows for easier automation. This is at most a minor limitation because normal wall is also associated with a dark rim. Hence, the algorithm classifies both types of stable walls, thick caps and normal vessel, as a single category. On the other hand, unstable fibrous caps that are thinned, eroded, or ruptured are separated by the algorithm.

8.5.5 Detection of Dark Rim

For each point along the contour, the gradient along the normal to the contour at that point is calculated. The TOF image is Wiener filtered to remove noise before gradient calculation. The actual gradient used is calculated as an average in a small (3 pixel) neighborhood of the normal. The gradient calculation extends 2 pixels into the lumen and 5–6 pixels outside the lumen (Fig. 8.26). This extent seems to cover most dark rims and also provide enough coverage to distinguish between a rim and a dark region next to the lumen.

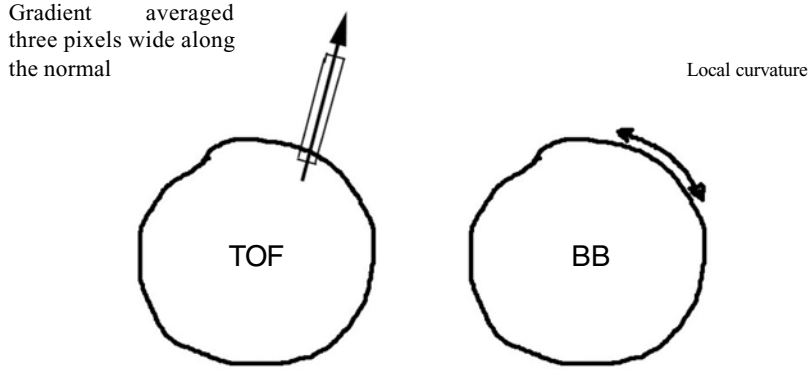


Figure 8.26: Feature vector calculation: Gradient along the normal to the TOF lumen contour and ratio of local curvature to global curvature on the registered BB lumen contour.

8.5.6 Detection of Focal Contour Abnormality

A focal contour abnormality is said to occur when the local curvature is large compared to the average lumen curvature. The curvature,

$$c = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} \tag{8.60}$$

is calculated for a small segment of the lumen and its ratio to the average curvature for the whole lumen is assigned to the point in the center of that segment. In order to obtain gradient and curvature parameters for the same point, the TOF contour and BB contour are brought into correspondence by registering the centroids of their convex hulls. With this definition any sharp change in curvature is detected. It becomes significant only when associated with the absence of a dark rim. However, it has to be noted that this could lead to some false classifications especially around the bifurcation.

8.5.7 Classification

The parameters for dark rim and focal abnormality were measured from a set of images identified by radiologists and confirmed by histology. Two sets each were used for thick, thin, and ruptured caps. Several measurements along the contour were thus available for each set. The mean and covariance of each parameter for thick, thin, and ruptured caps was then calculated. These templates were used for classification by the feature distance of a candidate point from a template for thick, thin, and ruptured classes. The Mahalanobis distance of the dark rim parameter was used to differentiate thick caps from the other two classes. The thin and ruptured classes were differentiated from the remaining points based on the curvature parameter again using the Mahalanobis distance metric,

$$r^2 = (x - m)C^{-1}(x - m) \tag{8.61}$$

where m and C are mean and covariance matrices, respectively. This decision is based on the observation that both thin and ruptured caps do not have a dark rim but the ruptured can be differentiated by the presence of a focal contour abnormality. Figure 8.27 shows an example of the algorithms classification compared to ground truth by histology.

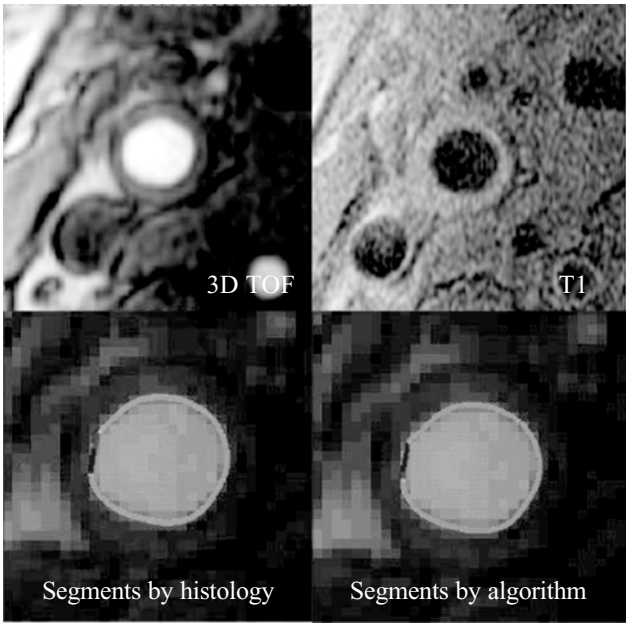


Figure 8.27: Example of FC classification with corresponding MR images and ground truth by histology. (A color version of this figure can be found on the CD. Green: thick cap; blue: thin cap; red: ruptured cap.)

8.5.8 Postprocessing

The classified pixels are then postprocessed to remove isolated classifications. These pixels are merged into those of surrounding pixels by a morphological opening operation (an element size of 5 was used). This makes the classification similar to what is outlined by a human operator so that classification by the algorithm can be compared to ground truth outlined by a pathologist.

8.5.9 Validation

A pathologist outlined the contour classifications from six endarterectomy patients. These sections were centered on the bifurcation with an average of 8–9 slices per patient. Fifty three sections out of these with matched MR image slices were chosen for analysis. The classification by the algorithm was then compared to the ground truth by histology. Each cap status per slice compared point by point for classification accuracy was used to calculate Pearson's correlation coefficients. The algorithm performs well in classifying thick and thin caps with

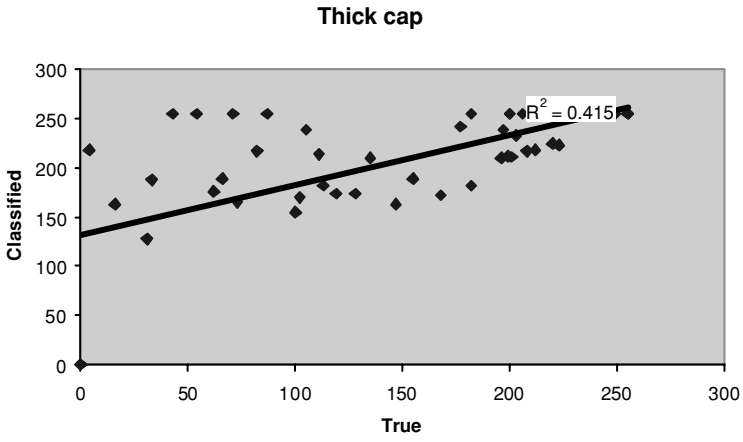


Figure 8.28: Correlation between true and classified pixels shows an $R = 0.6442$ for the thick cap (p value < 0.0001).

a correlation coefficient of 0.64 (significant with p value < 0.0001) and 0.62 (significant with p value < 0.0001) as shown in Figs. 8.28 and 8.29, respectively. The correlation coefficient for the ruptured cap is lesser (0.34, p value of 0.014) due to more false negatives and false positives. The correlation might be improved if specimen shrinkage [101] can be accounted for in matching correspondence between true and classified points. Differential shrinkage of the endarterectomy specimen during histological processing can cause twisting of the specimen around the arterial axis thus increasing classification error.

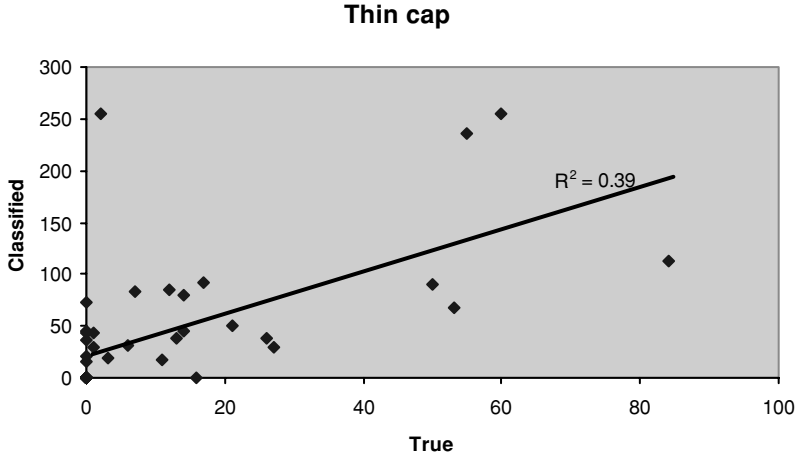


Figure 8.29: Correlation between true and classified pixels shows an $R = 0.6245$ for the thin cap (p value < 0.0001).

8.5.10 Conclusion

This preliminary algorithm shows promise in separating stable (thick) and unstable (thin) fibrous caps. Future work is aimed at improving the detection of ruptured cap and differentiating it from thin caps. Actual identification of ruptured caps is a more complicated issue involving multicontrast MRI with up to 5 weightings (3D TOF, T1, T2, PD, and contrast enhanced T1). A human expert also uses presence of juxtaluminal calcification, intraplaque hemorrhage, and thrombus to detect a ruptured cap. An algorithm that takes into account all the above weightings and factors would be more likely to differentiate ruptured caps from thin caps.

8.6 Conclusions

In this chapter, we discussed some postprocessing techniques to provide reliable and practical solutions for carotid plaque study based on MR images. For the three categories of images, single contrast weighting gray level images, image sequences, and multiple contrast weighting images, we have developed algorithms to address their specific needs and integrated into a software package, quantitative vascular analysis system (QVAS).

For single contrast weighting gray level image, we use MAP criterion with MRF priors as a powerful tool to build up the image model. The inherent noise-resistant ability and explicit description of pixel relations guarantee the results reliability and robustness. Also, the QHCF algorithm provides a feasible solution for its implementation in practical applications.

The solutions for image sequence segmentation and object tracking are built on the MRF-based active contour model. This framework incorporates the accurate and reliable region segmentation of MRF with the optimal contour tracking ability of minimal path approach. To ensure the optimal combination of these two models, a new criterion, maximum reliability, is set up as a bridge. This framework is also very flexible and extensible to include additional prior knowledge for various applications. In this study, it has been successfully applied to carotid artery tracking and lumen segmentation in MR image sequences.

Our initial study on multiple contrasts weighting MR image segmentation extends the MRF to multiple dimensions. However, because of the intrinsic limitations of this model, we adopted and further enhanced a clustering-based

algorithm by employing mean shift as density estimator. The results of multiple contrast weighting MR image segmentation and the histology section validation demonstrate very successful performance.

Detection of FC status is crucial for understanding the disease status and prognosis of atherosclerosis. The preliminary algorithm introduced in section 8.5 shows promise in separating stable (thick) and unstable (thin) FCs. Future work is aimed at improving the detection of ruptured cap and differentiating it from thin caps.

Since the images in our study are of poor quality than are usual practical images, the algorithms for gray level images and image sequences segmentation can be applied as general solutions. The multiple contrast weighting approaches can also be used for color images segmentation because their general properties are shared.

Questions

1. *What are the motivations and research directions in carotid artery atherosclerosis study?*
2. *Why is the study of constituents within carotid vessel wall very important?*
3. *Technically, what are the unique challenges in MRI obtained from advanced lesions in human carotid arteries?*
4. *What is the region segmentation method applied to single contrast MR image?*
5. *What is the advantage of using the MRF-based active contour model?*
6. *What are the criteria used in selecting the control points for active contour model?*
7. *How is dynamic weighting defined in multiple dimension MRF model?*
8. *What is dynamic mean shift density estimation in clustering multiple dimension data?*
9. *Why is automatic detection of fibrous cap status important?*
10. *What are the primary image features used in automatic fibrous cap detection?*

Bibliography

[1] Yuan, C., Mitsumori, L. M., Beach, K. W., and Maravilla, K.M., Carotid atherosclerotic plaque: Noninvasive MR characterization and identification of vulnerable lesions, Radiology, Vol. 221, No. 2, pp. 285–300, 2001.

[2] Savies, M. J. and Thomas, A. C., Plaque fissuring: The cause of acute myocardial infarction, sudden ischaemic death, and crescendo angina, Br. Heart J., Vol. 53, pp. 363–373, 1985.

[3] Falk, E., Stable versus unstable atherosclerosis: Clinical aspects, Am. Heart J., Vol. 138, No. 5(Pt.2), pp. 421–425, 1999.

[4] Davies, M. J., Richardson, P. D., Woolf, N., Katz, D. R., and Mann, J., Risk of thrombosis in human atherosclerotic plaques: Role of extracellular lipid, macrophage, and smooth muscle cell content, Br Heart J., Vol. 69, pp. 377–381, 1993.

[5] Fuster, V., Stein, B., Ambrose, J. A., Badimon, L., Badimon, J. J., and Chesebro, J. H., Atherosclerotic plaque rupture and thrombosis, evolving concepts, Circulation, Vol. 82, pp. 1147–1159, 1990.

[6] Kang, X. *et al.*, High resolution MRI of carotid atherosclerosis precision analysis of arterial lumen and wall area measurement, In: The 8th Scientific Meeting & Exhibition of the International Society for Magnetic Resonance in Medicine, Denver, CO, April 1–7, 2000.

[7] Yuan, C., Zhang, S., Polissar, N. L., Echelard, D., Ortiz, G., Davis, J. W., Ellington, E., Ferguson, M. S., and Hatsukami, T. S., Identification of fibrous cap rupture with magnetic resonance imaging is highly associated with recent transient ischemic attack or stroke, Circulation, Vol. 105, pp. 181–185, 2002.

[8] Toussaint *et al.*, MRI lipid, fibrous, calcified, hemorrhagic, and thrombotic components of human atherosclerosis in vivo, circulation, Vol. 94, pp. 932–938, 1996.

[9] Fu, K. S. and Mui, J. K., A survey on image segmentation, Patt. Recogn., Vol. 13, pp. 3–16, 1981.

[10] Haralic, R. M., and Shapiro, L. G., Image segmentation techniques, CVGIP: Graph. Models Image Process., Vol. 29, pp. 100–132, 1985.

[11] Beveridge, J. R. *et al.*, Segmenting image using localizing histograms and region merging, Int. J. Comput. Vision, Vol. 2, 1989.

[12] Leclerc, Y. G., Region growing using the MDL principle, In: DARPA Image Understanding Workshop, 1990.

[13] Trivedi, M. and Bezdek, J. C., Low-level segmentation of aerial images with fuzzy clustering, IEEE Trans. on System Man Cybern, Vol. 16, No. 4, pp. 589–598, 1986.

[14] Cpong, T., Shapiro, L. G., Watson, L. T., and Haralick, R. M., Experiments in segmentation using a facet model region grower, Comput. Vision Graph. Image Process, Vol. 1, pp. 360–372, 1972.

[15] Canny, J., A computational approach to edge detection, IEEE Trans. PAMI, Vol. PAMI-8, No. 6, pp. 679–698, 1986.

[16] Zhou, Y. T., Venkateswar, V., and Chellappa, R., Edge detection and linear feature extraction using a 2-D random field model, IEEE Trans. PAMI, Vol. 11, pp. 84–95, 1989.

[17] Haralick, R. M., Digital step edges from zero crossing of second directional derivatives, IEEE Trans. PAMI, Vol. 6, pp. 58–68, 1984.

[18] Reichenbach, S. E., Park, S. K., and Gartenberg, R. A., Optimal, small kernels for edge detection, In: Proceedings of 10th ICPR, 1990, pp. 57–63.

[19] Meier, T., Ngan, K. N., and Crebbin, G., A robust Markovian segmentation based on highest confidence first, In: IEEE International Conference on Image Processing, Santa Barbara, Oct. 1997.

[20] Pappas, T. N., An adaptive clustering algorithm for image segmentation, IEEE Trans. Signal Process., Vol. 40, No. 4, pp. 901–914, 1992.

[21] Chou, P. B. and Brown, C. M., The theory and practice of Bayesian image labeling, Int. J. Comput. Vision, Vol. 4, pp. 185–210, 1990.

[22] Geman, S. and Geman, D., Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. PAMI, Vol. PAMI-6, No. 6, pp. 721–741, 1984.

[23] Cohen, L. D. and Kimmel, R., Global minimum for active contour model: A minimal path approach, Int. J. Comput. Vision, Vol. 24, pp. 57–78, 1997.

[24] Wang, H. and Ghosh, B. K., Geometric deformable model and segmentation, In: IEEE International Conference on Image Processing, Chicago, USA, 1998.

[25] Vieren, C., Cabestaing, F., and Postaire, J. G., Catching motion objects with snakes for moving tracking, Patt. Recogn. Lett., Vol. 16, pp. 679–685, 1995.

[26] Bertalmio, M., Sapiro, G., and Randall, G., Morphing active contours: A geometric approach to topology-independent image segmentation and tracking, In: IEEE International Conference on Image Processing, Chicago, USA, 1998.

[27] Cohen, L. D., On active contour models and ballons, CVGIP: Image Understand., Vol. 53, No. 2, pp. 211–218, 1991.

[28] Bello, M. G., A combined Markov random field and wave-packet transform-based approach for image segmentation, IEEE Trans. Image Process., Vol. 3, No. 6, pp. 834–846, 1994.

[29] Zhu, S. C. and Yuille, A., Region competition: Unifying snake/balloon, region growing and Bayes/MDL/energy for multi-band image segmentation, In: Proceedings of Fifth International Conference on Computer Vision, 1995, pp. 416–423.

[30] Lumia, R., Shapiro, G., and Zuniga, O., A new connected component algorithm for virtual memory computers, Comput. Vision, Graphics Image Process., Vol. 22, pp. 287–300, 1983.

[31] Malladi, R., Sethian, J. A., and Vemuri, B. C., A topology independent shape modeling scheme, SPIE Geomet. Meth. Comput. Vision II, Vol. 2031, pp. 246–258, 1993.

[32] Lin, E., A Fuzzy Global Minimum Snake Model for Contour Detection, Ph.D. Dissertation, University of Washington, 1999.

[33] Besag, J., Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc. B, Vol. 36, No. 2, pp. 192–236, 1974.

[34] Yemez, Y., Sankur, B., and Anarim, E., Region growing motion segmenation and estimation in object-oriented video coding, ICIP, Vol. 2, pp. 521–524, 1996.

[35] Zhang, J. and Gao, J., Image sequence segmentation using curve evolution, In: 33th Annual Asilomar Conference on Signals, Systems and Computers, Oct. 1999.

[36] Wilson, R., Meulemans, P., Calway, A., and Krüger, S., Image sequence analysis and segmentation using G-blobs, ICIP, 1998.

[37] Alatin, A. A., Onural, L., Wollborn, M., Mech, R., Tuncel, E., and Sikora, T., Image sequence analysis for emerging interactive multimedia services—The European cost 211 framework, IEEE Trans. CSVT, Vol. 8, No. 7, pp. 802–813, 1998.

[38] Allen, J. T. and Huntsberger, T., Comparing color edge detection and segmentation methods, In: Proceedings of IEEE Southeaster Conference, 1989, pp. 722–728.

[39] Jain, A. K., Fundamentals of Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[40] Priese, L. and Rehrmann, V., On hierarchical color segmentation and applications, In: Proceedings of CVPR, New York, USA, 15–17 June 1993, pp. 633–634.

[41] Taylor, R. I. and Lewis, P. H., Color image segmentation using boundary relaxation, In: Proceedings of 11th IAPR International Conference on Pattern Recognition, Den Hague, Netherlands, Aug 30–Sept 2, 1992, Vol. III, pp. 721–724.

[42] Schettini, R., A segmentation algorithm for color images, Patt. Recogn. Lett., Vol. 14, pp. 499–506, 1993.

[43] Bonsiepen, L. and Coy, W., Stable segmentation using color information, In: Proceedings of CAIP'91, Dresden, Sept 17–19, 1991, pp. 77–84.

[44] Ferri, F. and Vidal, E., Color image segmentation and labeling through multiedit-condensing, Patt. Recogn. Lett., Vol. 13, No. 8, pp. 561–568, 1992.

[45] Umbaugh, S. E., Moss, R. H., Stoecker, W. V., and Hance, G. A., Automatic color segmentation algorithms with applications to skin tumor feature identification, IEEE Eng. Med. Biol., Vol. 12, No. 3, pp. 75–82, 1993.

[46] Chang, M. M., Patti, A. J., Sezan, M. I., and Tekalp, A. M., Adaptive Bayesian approach for color image segmentation, In: SPIE Conference on Visual Communication and Image Processing, Boston, MA, Nov 1993.

[47] Wright, W. A., Markov random field approach to data fusion and color segmentation, Image vision comput., Vol. 7, pp. 144–150, 1989.

[48] Comaniciu, D. and Meer, P., Robust analysis of feature space: color image segmentation, In: IEEE Conference Computer Vision and Pattern Recognition, Puerto Rico, 1997, pp. 750–755.

[49] Taxt, T., Flynn, P. J., and Jain, A. K., Segmentation of document images, IEEE Trans. PAMI., Vol. 11, No. 12, pp. 1322–1329, 1989.

[50] Nakagawa, Y. and Rosenfeld, A., Some experiments on variable thresholding, Patt. Recogn., Vol. 11, pp. 191–204, 1979.

[51] Weszka, J. S. and Rosenfeld, A., Threshold evaluation techniques, IEEE Trans. Syst. Man Cybern., Vol. 8, pp. 622–629, 1978.

[52] Pal, S. K. and Pal, N. R., Segmentation based on measures of contrast, homogeneity, and region size, IEEE Trans. Syst. Man Cybern., Vol. 17, pp. 857–868, 1987.

[53] Meyer, F. and Beucher, S., Morphological segmentation, J. Vis. Commun. Image Represent., Vol. 1, No. 1, pp. 21–46, 1990.

[54] Salembier, P. and Pardas, M., Hierarchical morphological segmentation for image sequence coding, IEEE Trans. Image Process., Vol. 3, No. 5, pp. 639–651, 1994.

[55] Li, S. Z., Markov Random Field Modeling in Computer Vision, Springer-Verlag, Berlin, 1995.

[56] Cerny, V., A thermo dynamical approach to the traveling salesman problem: An efficient simulation algorithm, J. Optimization Theory Appl., Vol. 45, pp. 41–51, 1985.

[57] Kirkpatric, S., Gellate, C. D., and Vecchi, M., Optimization by simulated annealing, Science, Vol. 220, pp. 671–680, 1983.

[58] Murray, D. W., Kashko, A., and Buxton, B. F., An approach to the picture restoration algorithm of Gemen and Geman on an SIMD machine, Image Vision Comput., Vol. 4, pp. 133–142, 1986.

[59] Metropolis, N., Equations of state calculations by fast computational machine, J. Chem. Phys., Vol. 21, pp. 1087–1092, 1953.

[60] Mokhtari, M. and Bergevin, R., Multi-scale segmentation and approximation for significant description of 2D contours, In: IEEE Conference on Image Processing, 1997.

[61] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput. Vision, pp. 321–331, 1988.

[62] Chiou, G. I. and Hwang, J. N., A knowledge driven stochastic active contour model (KBS-SNAKE) for contour finding of distinct features, IEEE Trans. Image Process., Vol. 4, No. 10, pp. 1407–1416, 1995.

[63] Caselles, V., Catte, F., Coll, T., and Dibos, F., A geometric model for active contours, Numerische Mathematik, Vol. 66, pp. 1–31, 1993.

[64] Miller, J. V., Breen, D. E., Lorensen, W. E., O'Bara, R. M., and Wozny, M. J., Geometrically deformed models: A method to extract closed geometric models from volume data, Comput. Graph., Vol. 25, No. 4, pp. 217–226, 1991.

[65] Osher, S. and Sethian, J. A., Fronts propagating with curvature dependent speed: Algorithms based on Hamiltion–Jacobi formulation, J. Comput. Phy., Vol. 79, pp. 12–49, 1988.

[66] Sethian, J. A., Numerical algorithm for propagating interface: Hamilton–Jacobi equation and conservation laws, J. Diff. Geometry, Vol. 31, pp. 131–161, 1990.

[67] Caselles, V., Kimmel, R., and Shapiro, G., Geodesic active contours, In: Proceedings of the Fifth International Conference on Computer Vision, Boston, MA, 1995, pp. 694–699.

[68] Amini, A. A., Tehrani, S., and Weymouth, T. E., Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints, In: Second International Conference on Computer Vision, 1990, pp. 95–99.

[69] Geiger, D., Gupta, A., Costa, L. A., and Vlontzos, J., Dynamic programming for detecting, tracking, and matching deformable contours, IEEE Trans. Patt. Anal. Machine Intell., Vol. 17, No. 3, pp. 294–302, 1995.

[70] Chandran, S. and Potty, A. K., Energy minimization of contours using boundary conditions, IEEE Trans. Patt. Anal. Machine Intell., Vol. 20, No. 5, pp. 546–549, 1998.

[71] Kass, M. *et al.*, Snakes: Active contour models, Int. J. Comput. Vision, pp. 321–331, 1987.

[72] Berger, M. O. and Mohr, R., Towards, autonomy in active contour models, In: Proceedings of 10th International Conference on Pattern Recognition, Atlantic City, NJ, USA, June 1990, Vol. 1, pp. 847–851.

[73] Yuan, C., Lin, E., and Hwang, J. N., Closed contour edge detection of blood vessel lumen and outer wall boundaries in black-blood MR images, Magn. Reson. Imaging, Vol. 17, No. 2, pp. 257–266, 1999.

[74] Cohen, L. D., On active contour models and balloons, CVGIP: Image Understand., Vol. 53, No. 2, pp. 211–218, 1991.

[75] Cohen, L. D. and Cohen, I., Finite-element methods, for active contour models and balloons for 2-D and 3-D images, IEEE Trans. Patt. Anal. Machine Intell., Vol. 15, No. 11, pp. 1131–1147, 1993.

[76] Lin, E., Hwang, J.-N., and Yuan, C., Measurements of blood vessel wall areas in black-blood MR images using global minimum snake algorithm, In: IEEE International Conference on Acoustic, Speech and Signal Processing, Phoenix, AZ, March 1999, Vol. 6, pp. 3409–3412.

[77] Yokoyama, T., Yagi, Y., and Yachida, M., Active contour model for extracting human faces, In: Fourteenth International Conference on Pattern Recognition, Brisbane, Qld., Australia, Aug 1998, Vol. 1, pp. 673–676.

[78] Wyszecki, G. and Stiles, W. S., Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd edn., Wiley, New York, pp. 113, 1982.

[79] Hunt, R. W. G., Measuring Color, Ellis Horwood, Chichester, England, 1987.

[80] Skarbek, W. and Koschan, A., Color image segmentation—A survey, Technical Report 94-32, Technical University, Berlin, October 1994.

[81] Clarke, S. *et al.*, Multispectral analysis of MR images of atherosclerotic plaque: Correlation with histology, In: 8th Annual Meeting of ISMRM, Denver, USA, April 2000.

[82] Panjwani, D. K., and Healey, G., Markov random field models for unsupervised segmentation of textured color images, IEEE Trans. PAMI, Vol. 17, No. 10, pp. 939–954, 1995.

[83] Comaniciu, D. and Meer, P., Mean shift analysis and applications, In: IEEE International Conference Computer Vision (ICCV'99), Kerkyra, Greece, 1999, pp. 1197–1203.

[84] Pal, N. R. and Pal. S. K., A review on image segmentation techniques, Patt. Recogn., Vol. 26, No. 9, pp. 1277–1294, 1995.

[85] Rousseeuw, P. J. and Leroy, A., Robust Regression and Outlier Detection, Wiley, New York, Section 7.1, 1987.

[86] Jain, A. K., Murty, M., Narasimha, and Flynn, P. J., Data clustering: A review, ACM Comput. Surv., Vol. 31, No. 3, pp. 264–323, 1999.

[87] Rousseeuw, P. J., Least median of squares regression, J. Am. Stat. Assoc., Vol. 79, pp. 871–880, 1984.

[88] Jolion, J. M., Meer, P., and Bataouche, S., Robust clustering with applications in computer vision, IEEE Trans. Patt. Anal. Machine Intell., Vol. 13, pp. 791–802, 1991.

[89] Fukunaga, K. and Hostetler, L. D., The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Trans. Info. Theory, Vol. IT-21, pp. 32–40, 1975.

[90] Cheng, Y., Mean shift, mode seeking, and clustering, IEEE Trans. Patt. Anal. Machine Intell., Vol. 17, pp. 790–799, 1995.

[91] Silverman, B. W., Density Estimation for Statistics and Data Analysis, Chapman and Hall, New York, 1986.

[92] Xu, D. and Hwang, J.-N., A topology independent active contour tracking, In: IEEE NNSP'99, Madison, USA, Aug 23–25, 1999, pp. 164–167.

[93] Wasserman, B. A., Clinical carotid atherosclerosis, Neuroimaging Clin. N. Am., pp. 403, 2002.

[94] Willeit, J. and Kiechl, S., Biology of arterial atheroma, Cerebrovas. Dis., Vol. 10, pp. 1–8, 2000.

[95] Dhume, A. S., Soundararajan, K., Hunter, W. J., and Agrawal, D. K., Comparison of vascular smooth muscle cell apoptosis and fibrous cap morphology in symptomatic and asymptomatic carotid artery disease, Ann. Vas. Surg., Vol. 17, pp. 1–8, 2003.

[96] Winn, W. B., Schmiedl, U. P., Reichenbach, D. D., Beach, K. W., Nghiem, H., Dimas, C., Daniel, E., Maravilla, K. R., and Yuan, C., Detection and characterization of atherosclerotic fibrous caps with T2-weighted MR, Am. J. Neuroradiol., Vol. 19, pp. 129–134, 1998.

[97] Hatsukami, T. S., Ross, R., Polissar, N. L., and Yuan, C., Visualization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque in-vivo with high resolution magnetic resonance imaging, Circulation, Vol. 102, pp. 959–964, 2000.

[98] Mitsumori, L. M., Hatsukami, T. S., Ferguson, M. S., Kerwin, W. S., Cai, J. C., and Yuan, C., In vivo accuracy of multisequence MR imaging for identifying unstable fibrous caps in advanced human carotid plaques, J. Magn. Reson. Imaging, Vol. 17, pp. 410–420, 2003.

[99] Wasserman, B. A., Smith, W. I., Trout, H. H., Cannon, R. O., Balaban, R. S., and Arai, A. E., Carotid artery atherosclerosis: In vivo morphologic characterization with gadolinium-enhanced double-oblique MR imaging—initial results, Radiology, Vol. 223, pp. 566–573, 2002.

[100] Han, C., Hwang, J. N., and Yuan, C., A fast minimal path active contour model, IEEE Trans. Image Process., Vol. 6, pp. 865–873, 2001.

[101] Eubank, W. B., Yuan, C., *et al.*, Endarterectomy specimen shrinkage: Comparison of T2-weighted MR imaging of specimen ex vivo to histological process, J. Vascular Invest., Vol. 4, pp. 147–152, 1998.

Chapter 9

Accurate Lumen Identification, Detection, and Quantification in MR Plaque Volumes

Jasjit Suri,[1] *Vasanth Pappu,*[1] *Olivier Salvado,*[1] *Baowei Fei,*[1] *Swamy Laxminarayan,*[2] *Shaoxiong Zhang,*[3] *Jonathan Lewin,*[3] *Jeffrey Duerk,*[3] *and David Wilson*[1]

9.1 Introduction

The importance of plaque component classification and vessel wall quantification has been well established by several research groups (see Refs. [1–30]). Following are the two main reasons for this research:

1. *Regression and progression of atherosclerosis:* Direct plaque imaging is of potential use not only for diagnosis but also for monitoring response to treatment. Angiographic studies of progression and regression of atherosclerosis have been notoriously poor in demonstrating changes in plaque burden, even when changes in clinical event rates have been markedly altered (see Brown *et al.* [6]). In a study of diet-/injury-induced atherosclerosis in rabbits, T_2-weighted MRI identified regression of atherosclerosis 12–20 months after the withdrawal of the atherogenic diet (regression group). In contrast, lesion progression was documented in rabbits that were continued on the atherogenic diet (progression group). Helft *et al.* [25] showed that there was a significant reduction in the lipid

[1]Biomedical Engineering Department, Case Western Reserve University, Cleveland OH, USA
[2]Biomedical Engineering Department, Idaho State University, Pocatello, ID, USA
[3]Department of Radiology, Case Western Reserve University, Cleveland OH, USA

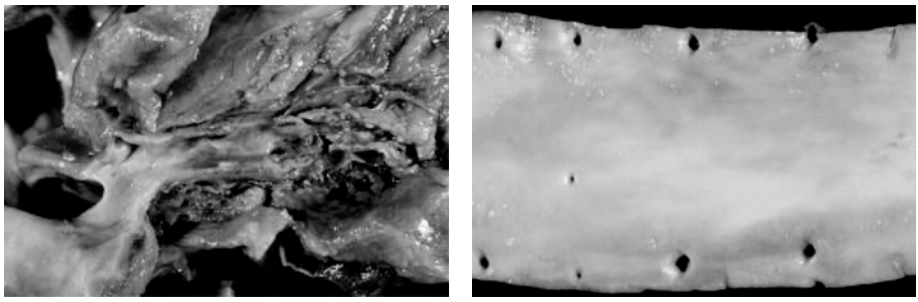


Figure 9.1: Cross section of the artery showing lipids.

components of the plaque in the regression group and an increase in the progression group.

2. *Importance of wall area:* In a preliminary analysis, using PD-weighted and T_2-weighted MRI, Corti *et al.* [27] illustrated a decrease in cross-sectional wall area in atherosclerotic segments of human aorta and carotid artery (by 8% and 15%, respectively) 12 months after the initiation of simvastatin. More importantly, there was no change in cross-sectional area of the arterial lumen. This emphasizes the importance of imaging the vessel wall directly and probably explains the limitations of coronary angiography in assessing response to treatment.

9.1.1 Histological Description of the Lumen

Figure 9.1 shows the 3-D view of the cross section of the lumen. Figure 9.2 shows the histological cross section of the artery. There are three layers in the walls of both arteries and veins:

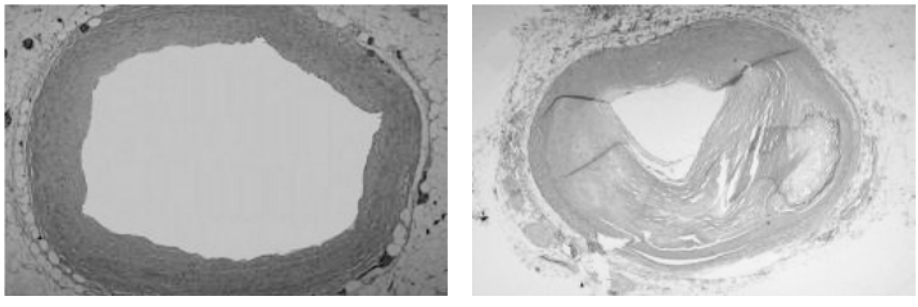


Figure 9.2: Histology image of the arterial cross section.

1. *Tunica intima:* It is the innermost layer which consists of the endothelium (a simple squamous epithelium) and a small amount of underlying connective tissue. In arteries it also includes the internal elastic lamina, which is often seen as a thick wavy band surrounding the lumen of the vessel.

2. *Tunica media:* This is the middle layer made up mainly of smooth muscle cells. In arteries it is the thickest layer.

3. *Tunica adventitia:* It is the outer layer made up of loose connective tissue (collagen fibers, fibroblasts) along with some smooth muscle cells. It is the thickest layer in veins (particularly the larger veins).

Thus we see that quantification of walls and the classification of plaque components is of utmost importance. In the next section we discuss the research groups who have done work in this direction.

9.1.2 Survey of Plaque Segmentation Techniques

Figure 9.3 shows the different image processing techniques used for segmentation of the plaque volumes. Yuan *et al.* [33] used a quantitative vascular analysis tool (QVAT). The QVAT is a semiautomatic, custom-designed program that tracks boundaries and computes areas. Gill *et al.* [9] used a mesh-based model that obtained boundaries in three steps. It involved a deformable balloon model of a triangular mesh which is first placed inside a region manually; it is then inflated by inflation forces and then refined by image-based forces. Kim *et al.* [11] used an edge-detection tool. Wilhjelm *et al.* [12] used a manual segmentation procedure. Yang *et al.* [10] used a border-based model, which had three steps. It involved first approximating the outlines of the vessels, followed by the detection of borders, and then the user correction of the borders.

Yang *et al.* [10] segmented the wall and plaque in *in vitro* vascular MR images using a combination of automated and manual processes. A computerized method used edge strength, edge direction, border smoothness, and shape guidance to get the outer wall, lumen, internal elastic lamina, and external elastic lamina boundaries; these boundaries were then modified using the user selection seed points. Boundaries of the outer wall and the lumen were determined by fitting splines to seed points. The internal elastic lamina boundary was obtained

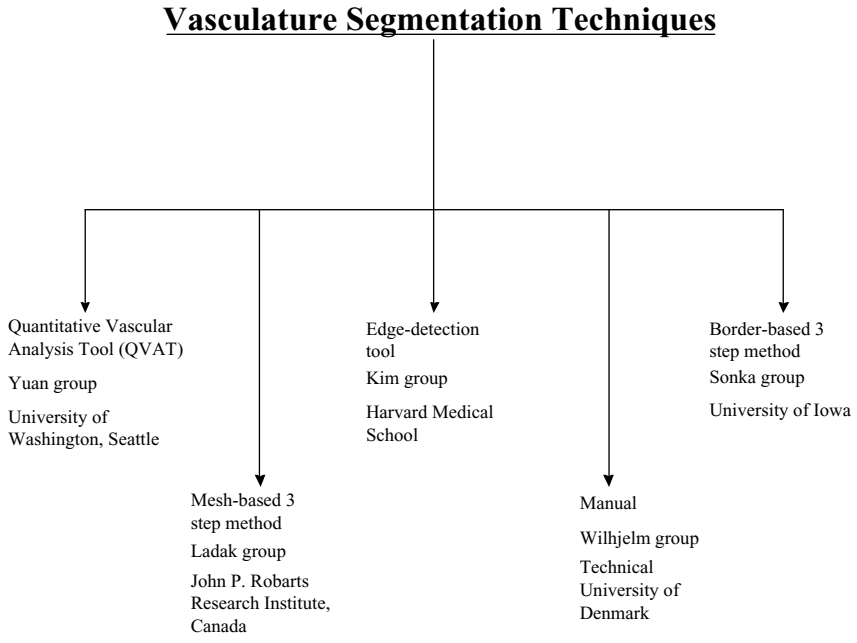


Figure 9.3: Segmentation techniques applied to plaque imaging volumes.

using the lumen wall boundary for shape guidance. The external elastic lamina boundary was then obtained using the internal elastic lamina for shape guidance.

Kim *et al.* [11] imaged the proximal coronary artery vessel wall with high-resolution 3-D cardiovascular MRI. The proximal vessel wall and lumen boundaries were obtained using an automated edge detection tool. Comparison of vessel wall thickness and luminal diameter between healthy subjects and patients with coronary artery disease showed increased wall thickness and but no significant difference in luminal diameter in patients. This was due to positive arterial remodeling in the patients, known as the *Glagov effect.*

Wilhjelm *et al.* [12] spatially compounded ultrasound images of *formalin*-fixed carotid atherosclerotic plaques to reduce angle-dependence and speckle noise—two problems prominent in ultrasound. A digital off-line ultrasound scanner for multi-angle compound imaging (MACI) produced arterial image slices that were compared to the corresponding anatomical slices. Compared to B-mode ultrasound images, the MACI images had a better definition of outlines and a more uniform representation of tissue parameters, which can aid in the diagnosis of atherosclerotic disease.

Jespersen *et al.* [13] compared ultrasound B-mode images and histological analysis of carotid plaque. Patients with carotid disease were scanned before carotid endarterectomy was performed. The removed plaques were then fixed and histologically analyzed. The ultrasound images were video recorded and the plaques were outlined with the help of color flow mapping. A frame grabber was used to convert the video recorded scannings into digital 256-gray level images stored on computer. Text features were calculated from the digital images, one of which was the gray level co-occurrence matrix (GLCM). Second-order text features derived from the GLCM were used to classify the plaques into three constituents, soft materials, fibrous tissue, and calcification, where soft materials included lipid, blood, and thrombus. These features were found to correlate well with the visual analysis; however, both classifications did not correlate strongly with the histological analysis because of the *echolucent plaques* of B-mode imaging.

Quick *et al.* [14] reviewed the concepts of MR imaging of the vessel wall. MRI can be used to determine the nature of the atherosclerotic plaques; classification is based on signal intensities and morphological appearance in the different MR imaging modalities. Noninvasive phased-array radio-frequency coils are used to resolve the trade-off problem between high SNR and signal penetration depth in choosing RF coils. Imaging of the carotid and the coronary arteries has been discussed. Intravascular receiver coils offer better image quality and resolution, but have drawbacks such as being noninvasive, having wall motion artifacts, and occluding bloodflow. Stents can function as receiving coils by either connecting a cable to the stent, or inducing the stent from outside. Ultrasmall superparamagnetic particles of iron oxides (USPIOs) can be used as a contrast agent for detecting atherosclerotic plaques before luminal narrowing because of the susceptibility-induced signal voids they cause after being *phagocytosed by macrophages*.

Gill *et al.* [9] developed a semiautomatic segmentation technique based on an inflating model that they used to segment the lumen from three-dimensional ultrasound images of the carotid arteries of phantoms and subjects. The vascular mimicking phantom consisted of two vessels that were identical except that one of them was cut to simulate an *ulceration*. After the phantom was imaged, the two vessels were *registered* using the automatic nonlinear image matching and anatomical labeling algorithm. The segmentation algorithm involved interactive placement of the initial balloon model inside the lumen, automatic inflation of

the balloon using inflation forces, and automatic localization of the balloon to the arterial wall using image-based forces. The balloon model was represented by a triangular mesh. Two thresholds were used in defining when a triangle in the mesh should be split into two triangles; the larger threshold was used while the balloon was inflating to the arterial wall, and the smaller threshold was used while the balloon was being refined to fit the arterial wall. Surface tension was used to reduce the effect of noise. A maximum error corresponding to the maximum separation of the two registered phantom arteries was reported to be 0.3 mm.

Zhang *et al.* [35] showed that images produced by different imaging modalities of MRI will give similar results when measuring lumen and vessel wall areas, provided that the quality of the images are high and comparable. An image quality rating criteria was developed and had five levels of quality. Ten patients were imaged with four MRI modalities (Time of Flight, T_1, T_2, PD-weighted), and image sets of a patient were studied only if all of the different images were above the third level of image quality. Lumen and outer wall boundaries were measured semiautomatically using a program called the quantitative vascular analysis tool (QVAT). Since flow artifacts were better suppressed on double inversion T_1-weighted images, those images were recommended for measurement when those images have the highest image quality. Mean differences between lumen area measurements of each of the three black blood imaging techniques were shown to be not statistically significant. In measurements of lumen area, outer wall boundary area, and wall area, the PD and T_2-weighted images showed the best agreement.

Yuan *et al.* [33] studied whether using a gadolinium-based contrast agent in high-resolution MRI provided additional information that helped in characterizing atherosclerotic plaques. High-risk atherosclerotic plaques were characterized by thinning and rupture of the fibrous cap overlying the thrombogenic lipid core of the artery. The study was done on patients scheduled for carotid endarterectomy and volunteers. High-resolution cross-sectional MR images of bilateral carotid arteries were obtained with a phased array carotid coil on a 1.5-T GE SIGNA Horizon Echo Speed 5.8 MR scanner using a pre- and postcontrast-enhanced double inversion recovery T_1-weighted fast spin-echo imaging protocol with TR/TE/T1 = 800/10/650 msec, echo train length = 8, slice thickness = 2 mm, FoV = 13×9 cm, and matrix = 512×512 with zero-filled Fourier reconstruction. TOF images were also obtained to aid in the classification of

plaque tissues. The precontrast enhanced images were used to identify regions of interests (ROIs) in which the constituents were classified as fibrous tissue, necrotic core, or calcification. These ROIs were then matched in the postcontrast enhanced images and a percent signal intensity change was calculated from each ROI. After the endarterectomy, the plaques were histologically classified. Results were analyzed using statistical techniques such as single-factor analysis of variance (ANOVA), Tukey, and Student's t test. It was found that the use of the gadolinium-based contrast agent in MRI is significantly useful in the classification of necrotic core, fibrotic tissue, and especially neovasculature of atherosclerotic plaques.

Yuan *et al.* [34] further showed that identification of a ruptured fibrous cap in *in-vivo* human carotid atherosclerosis using high-resolution MRI is highly associated with a recent transient ischemic attack (TIA) or stroke. Multiple contrast-weighted MR protocol was used to obtain the images. The fibrous caps were reviewed and classified as either being intact and thick, intact and thin, or ruptured. Patients were classified as symptomatic or asymptomatic depending on recent history of TIA or stroke. It was observed that while 9% of patients with thick fibrous caps were symptomatic, 50% and 70% of those which had thin caps and ruptured caps, respectively, were symptomatic. Statistical analysis showed a highly significant trend of increasing percent symptomatic as cap deterioration increases.

Naghavi *et al.* [22, 23] discussed a new classification system for identifying patients having a risk of cardiac disease and related events. They defined three areas of vulnerability: plaque, blood, and myocardium. They defined a vulnerable plaque with a set of major and minor criteria, and techniques for detection of each of these criteria. Many markers in blood that were associated with coagulation were described, as were conditions that are associated with a vulnerable myocardium. A new risk assessment strategy that was based on the three areas of vulnerability, called the Cumulative Vulnerability Index, was proposed.

Fayad *et al.* [24] discussed the use of electron-beam computed tomography (EBCT) to quantitatively detect the amount of calcium deposited in the coronary arteries. Using a multidetector-row CT (MDCT) system to detect calcium offers higher spatial resolution and SNR, but has more motion artifacts. Additionally, using a contrast agent with MDCT can classify plaques into soft, intermediate, or calcified. EBCT angiography results were found to be similar to MDCT angiography results. Coronary MR angiography (CMRA) was still less sensitive

and specific than EBCT and MDCT angiographies; both spatial and temporal resolution were lower, and the time needed to acquire an image requires that imaging take place over multiple heart beats. MRI has been shown to usefully image plaques at various locations. It has also been used to monitor experimental studies on plaque [25]. The combination of CT and MRI for use in detecting dangerous plaques was promising.

Corti *et al.* [27] used high-resolution MRI imaging to follow the effects of *simvastatin*, a statin that stabilizes plaques by lowering the lipid content, on human atherosclerotic plaques. Results showed that after 12 months there was a significant decrease in vessel wall area and maximal wall thickness, but there was not a significant change in the lumen area.

Fuster *et al.* [26] discussed the biological events that lead to acute coronary syndromes (ACS). Plaques of types IV and V (vulnerable) and type VI (complicated) were most likely to lead to ACS. The beginning of an atherosclerotic lesion start with lipoprotein transport and development of the extracellular matrix. The disruption of plaques was made up of passive and active phenomenon. Inflammatory cells at the plaque site would weaken the fibrous cap through lytic processes, which was a step in arterial remodeling. Tissue factor (TF) was associated with *macrophages* and was involved in *coagulation*, *haemostasis*, and *thrombosis*. It was recognized that MRI is a promising tool for noninvasive plaque characterization.

9.1.3 What Is This Chapter About?

The formation of atherosclerotic plaques in vessel walls cause stenosis and is a major cause of death in the United States. Quantification of the degree of stenosis can lead to life-saving surgery. We want to quantify the boundaries of the lumen to determine the degree of the stenosis. This information can be used to diagnose surgery. We want to use MRI images because of its high resolution and ability to delineate the lumen wall. The MRI images of human carotid arteries we studied had an outer boundary and an inner boundary. The inner boundary was the boundary of the lumen.

We discuss modeling the lumen region, lumen boundary, and lumen quantification in this chapter. Also discussed is the analysis and quantification of vessel walls. For the outer boundary, the research will appear elsewhere.

We developed a system which analyzed and quantified the inner boundary of the lumen. Given a slice of an MRI image of the left and right carotid arteries, the

system detects and identifies the two different left and right lumen boundaries and quantifies them. The lumen is complicated to classify, since the blood in the lumen flows parabolically. Blood in the center of the lumen flows at a higher speed than the blood near the edges of the lumen. In an MRI image this difference in flow rates causes the center of the lumen to appear brighter than the edges of the lumen. When classifying the image, the classifier will fail to identify the entire lumen as one class, instead it will identify multiple classes inside the lumen. We used three different segmentation methods for the classification of the lumen region in our system. These are the Markov random fields (MRF), the Fuzzy C means (FCM), and the graph segmentation methods (GSM). The MRF method uses the Bayes rule to segment the image. It uses the expectation-maximization (EM) algorithm and is based on maximum likelihood. It segments the image into a given number of classes. The FCM method is based on the clustering technique. It computes the fuzzy membership function. It associates this function to each pixel in image. The GSM method is based on analyzing the image as a graph with the pixels being nodes and the edges being the connections between two pixels. It calculates weights of the edges and decides with a decision criterion whether there should be a boundary between them. After the image is classified using one of the three methods of segmentation, the image is binarized to isolate the left and the right lumens. Since the lumens may contain multiple classes, the binarization process merges these classes when necessary. The carotid arteries bifurcate in the middle of the volume and the region of interest (ROI) of the lumens change from being a circular shape to being an elliptical shape. The binarization process uses both circular and elliptical masks. Once the boundaries of the left and right lumens were obtained, they were compared to traced ground truth boundaries using two methods of error computation between boundaries. We computed the error using the shortest distance method (SDM) and the polyline distance method (PDM). The PDM computes a lower error than does the SDM. We tested the system for the three different classifying methods, first on synthetic data and then on real patient volumes.

We created a model of images of the carotid arteries for validation. To simulate noise, we created images with variance from 0 to 100 for a small noise protocol, and we created images with variance from 100 to 1000 for a large noise protocol. For each variance we created images with left and right lumens having two classes in eight different orientations. Each protocol had about 24,000 total boundary points. Using MRF, the average error for a variance of 500 pixels squared was 5.97 pixels with standard deviation of

0.13 pixels; using FCM the average error was 1.54 pixels with standard deviation of 0.05 pixels.

We ran the system using each of the three different classifying methods on real patient data. Ground truth boundaries of the walls of the carotid artery were traced for 15 patients. Overall the number of boundary points was roughly 22,500 points. A pixel was equivalent to 0.25 mm. Using MRF, the average error was **0.61** *pixels; using FCM, the average error was* **0.62** *pixels; using GSM, the average error was* **0.74** *pixels.*

What is new in this chapter? The following are the new things the readers will observe when it comes to plaque imaging: (a) Application of three different sets of classifiers for lumen region classification in plaque MR protocols. These classifiers are done in multiresolution framework. Thus subregions are chosen and subclassifiers are applied to compute the accuracy of the pixel values belonging to a class. (b) Region merging for subclasses in lumen region to compute accurate lumen region and lumen boundary in cross-sectional images. (c) Rotational effect of ROI in bifurcation zones for accurate lumen region identification and boundary estimation.

9.2 Challenges in Lumen Wall Boundary Estimation

Following are the challenges for lumen wall (inner) and vessel wall (outer) estimation processes (see also Fig. 9.4):

1. *Mulitple classes in the lumen region due to laminar blood flow:* The lumen region consists of multiple classes: core class (central part of the lumen), adjoining class (due to slow moving blood flow as seen in Fig. 9.4), and some times border pixels in the fibrous cap region giving different classes. So, the lumen region can be **C1**, **C1** + **C2**, or **C1** + **C2** + **C3** class regions.
2. *Lumen shape variation:* The shape of the cross section of the artery lumen is "circular" for some slices and is "elliptical" near the bifurcation. So, the ROI can change from slice to slice also. If one uses a circular ROI on an elliptical region, then a large number of pixels will be missed along the major axis of the elliptical region. The elliptical region can be seen on slice before the bifurcation zone, while the circular regions can be seen on slices far from the bifurcation zone.

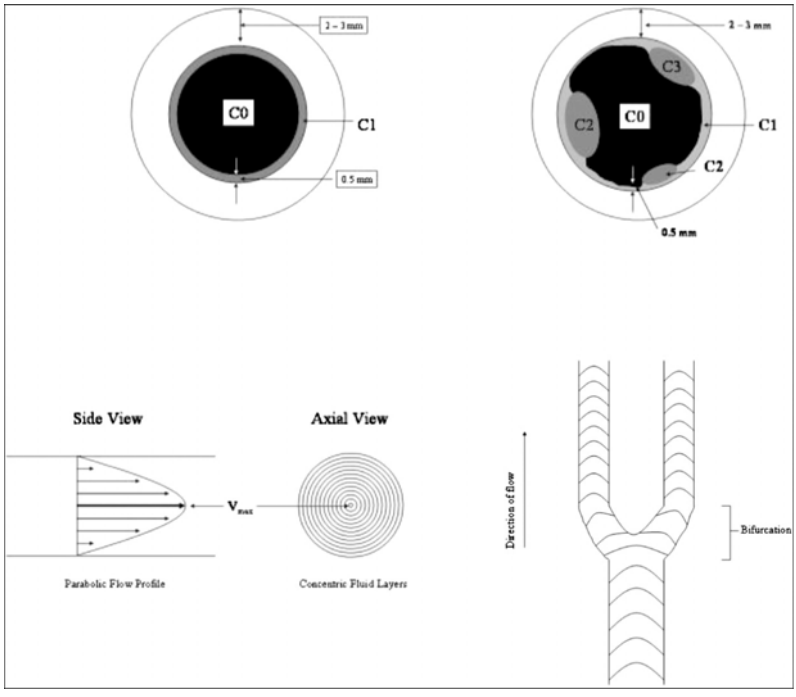


Figure 9.4: Top left: Figure showing the class **C0** for lumen, **C1** class for low intensity flow region, and the outer wall for the vessel wall. Top right: Classes **C1**, **C2**, and **C3** are the regions due to the classification process due to weak distribution of pixels in the boundary region. Bottom left: Parabolic flow of the blood showing the highest velocity in the central region of the vessel. Bottom right: Flow of blood in the bifurcation zone.

3. *Over-shooting of the human tracings:* Another difficulty which can bring large error is when the human tracing the ideal boundary overshoots the lumen region and draws in the vessel wall area or even outside of the vessel wall area. This "overshoot" tracing can bring large error between the computer estimated boundary and "ideal boundary."

4. *Bleeding region of the lumen:* The bleeding issue is a serious problem. Sometimes lumen class **C1** or **C2** or **C3** are not isolated. These class regions tunnel into the neighboring region and bleed, creating a break in the wall boundary or a missing boundary region.

5. *Partial volume effect:* The partial volume effect in the edge of the lumen can lead to misleading lumen wall boundary estimation.

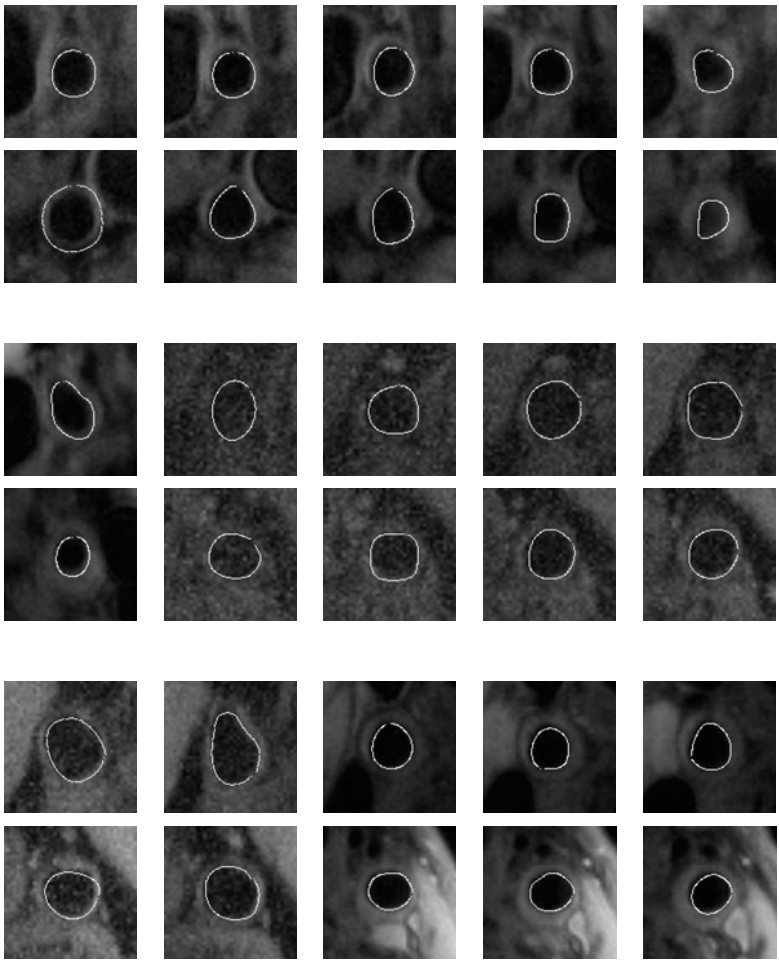


Figure 9.5: Abnormal ground truth overlays. The top row in each row pair is the left carotid artery overlayed with the ground truth tracing of the inner lumen wall, and the bottom row is the corresponding image of the right carotid artery. Some lumens have a circular shape, while others have an elliptical shape.

9.2.1 Ground Truth Tracing and Data Collection

Figures 9.5–9.9 show abnormal images of the left and right carotid arteries overlayed with the ground truth tracing. In each pair of rows, the top row is the left carotid arteries and the bottom row is the right carotid arteries. The ground truth tracing is the boundary of the inner lumen wall. Figure 9.9 shows the normal images of the same left and right carotid arteries. The normal images show

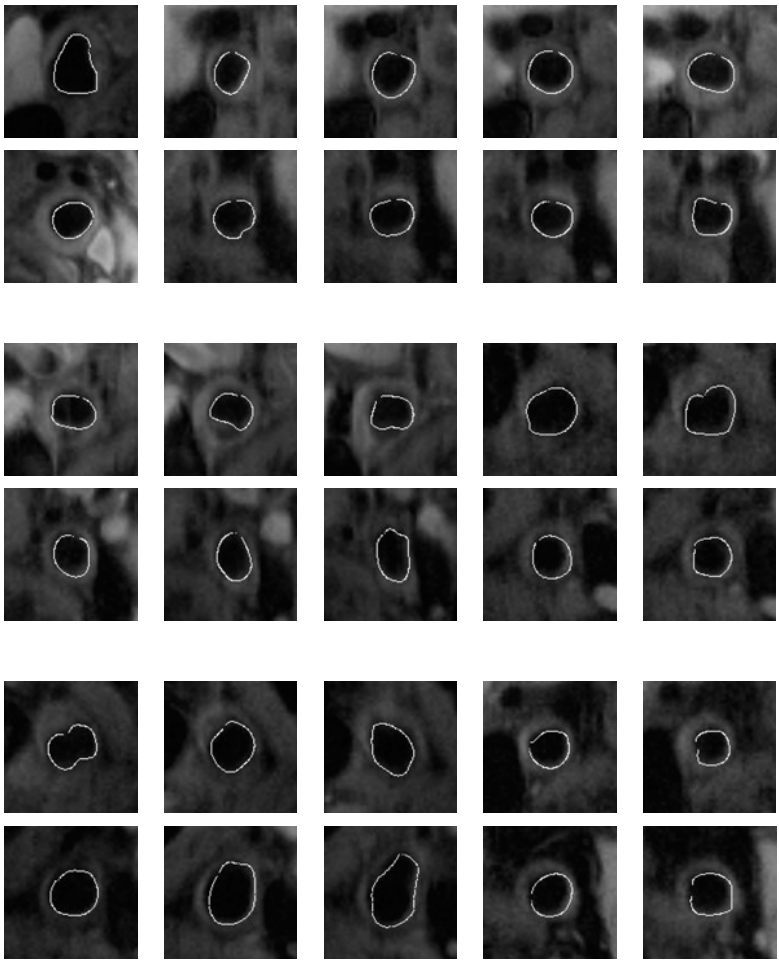


Figure 9.6: Abnormal ground truth overlays. The top row in each row pair is the left carotid artery overlayed with the ground truth tracing of the inner lumen wall, and the bottom row is the corresponding image of the right carotid artery. Some lumens have a circular shape, while others have an elliptical shape.

lumens that are more circular than those of the abnormal images, and there is no constriction of the arteries. Ground truth tracing was done using the MATLAB program MRI GUI ver 1.2. For each lumen boundary the image was zoomed in and points were plotted by the user around the inner wall boundary. The points were spline-fitted to 20 points.

The imaging parameters (TR/TE/TI/NEX/thickness/FOV/ETL) are as following: T1W:1R-R/7.1 ms/500 ms/2/3 mm/12–14 cm/21; PDW: 2R-R/7.1 ms/600 ms/

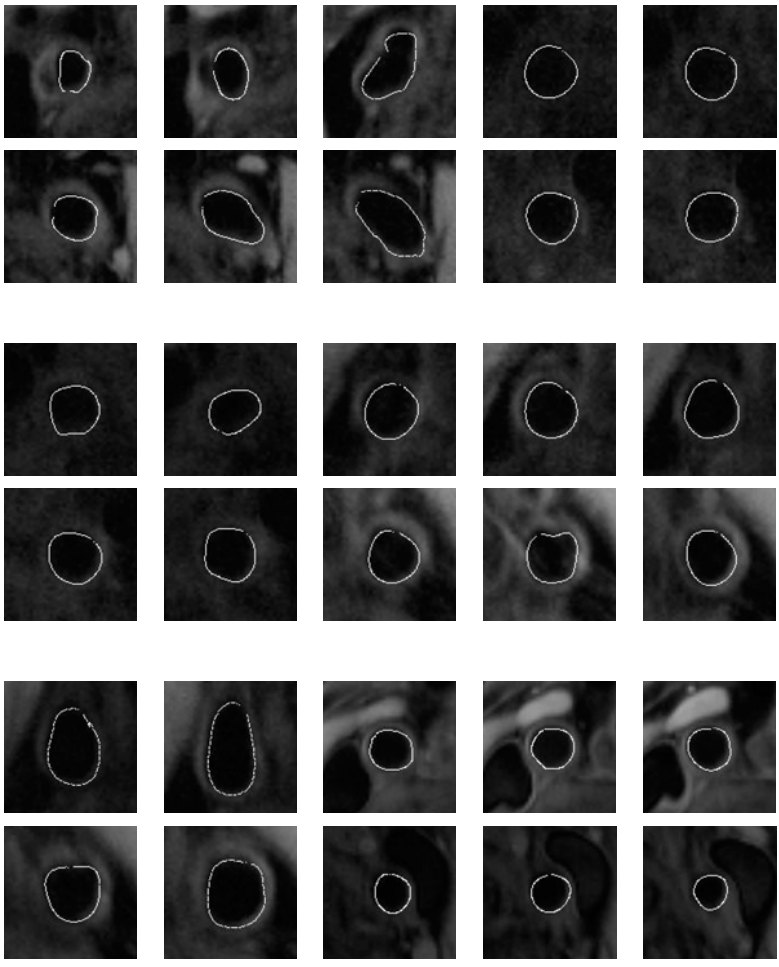


Figure 9.7: Abnormal ground truth overlays. The top row in each row pair is the left carotid artery overlayed with the ground truth tracing of the inner lumen wall, and the bottom row is the corresponding image of the right carotid artery. Some lumens have a circular shape, while others have an elliptical shape.

2/3 mm/12–14 cm/31; T2W: 2R-R/68 ms/600 ms/2/3 mm/12–14 cm/31. Matrix for all images were 256×192. Voxel size was $0.5 \times 0.5 \times 3$ mm^3. For bright blood 3D TOF images: TR/TE/flip angle/thickness/FOV = 20 ms/3.4 ms/25/2.0 mm/18. Zero-filled Fourier transform was used to create voxels of $0.35 \times 0.35 \times 1$ mm^3 for 3-D TOF imaging. The factors which affect MR plaque quality are (a) random patient motion, including (b) obese patients who may have deeper carotid arteries, (c) incomplete flow suppression, and (d) artery wall pulsation.

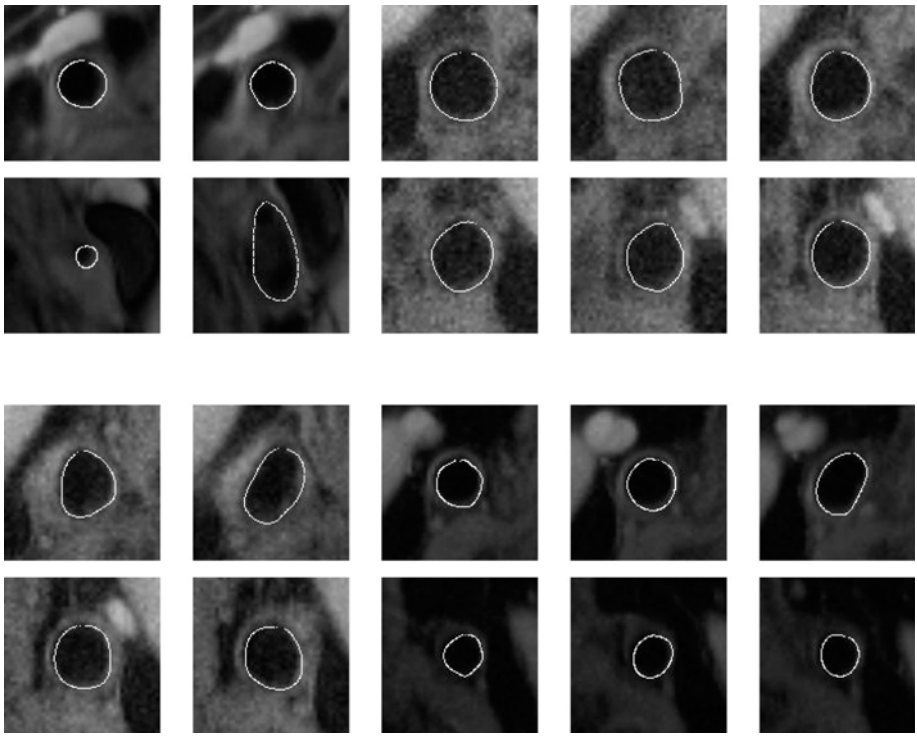


Figure 9.8: Abnormal ground truth overlays. The top row in each row pair is the left carotid artery overlayed with the ground truth tracing of the inner lumen wall, and the bottom row is the corresponding image of the right carotid artery. Some lumens have a circular shape, while others have an elliptical shape.

9.3 Three Pixel Classification Algorithms: MRF, FCM, and GSM

9.3.1 Markovian-Based Segmentation Method

The algorithm consisted of running the pixel classification approach using Markov random field with mean field (see Zhang [149]). Here, the image segmentation was posed as a classification problem where each pixel is assigned to one of K image classes. Suppose the input image was $\mathbf{y} = \{\mathbf{y}_{i,j}, (i, j) \in \mathbf{L}\}$, where $\mathbf{y}_{i,j}$ is a pixel, i.e., a 3-D vector, and $\mathbf{L}$ was a square lattice. Denote the segmentation as $\mathbf{z} = \{\mathbf{z}_{i,j}, (i, j) \in \mathbf{L}\}$. Here, $\mathbf{z}_{i,j}$ is a binary *indicator vector* of dimension K, with only one component being 1 and the others being 0. For

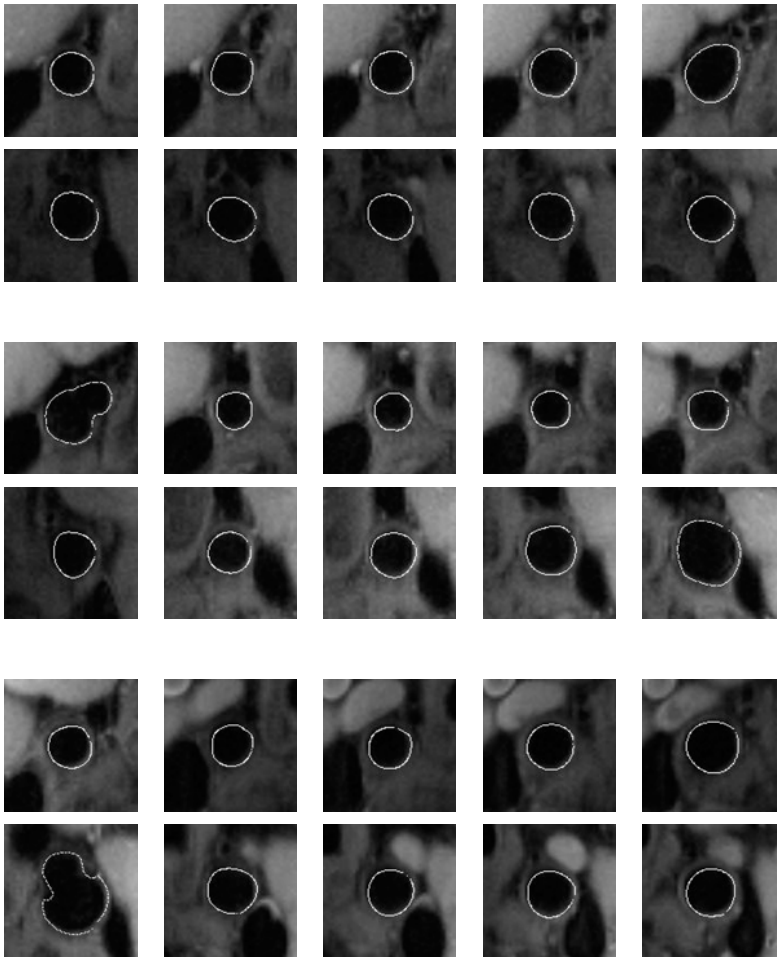


Figure 9.9: Normal ground truth overlays. The top row in each row pair is the left carotid artery overlayed with the ground truth tracing of the inner lumen wall, and the bottom row is the corresponding image of the right carotid artery. Some lumens have a circular shape, while others have an elliptical shape.

example, when $K = 3$, $\mathbf{z}_{i,j} = [0, 1, 0]^T$ means we assign the pixel at (i, j) to class 2.

Using the notation introduced above, the segmentation problem can be formulated as the following MAP (maximum a posteriori) inference problem:

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} [\log p(\mathbf{y} \mid \mathbf{z}, \Phi) + \log p(\mathbf{z} \mid \Theta)], \tag{9.1}$$

where Φ and Θ were model parameters. In this work, we assume that the pixels in $\mathbf{y}$ are conditionally independent given $\mathbf{z}$, i.e.,

$$\log p(\mathbf{y} \mid \mathbf{z}, \Phi) = \sum_{i,j} \log p(\mathbf{y}_{i,j} \mid \mathbf{z}_{i,j}, \Phi). \tag{9.2}$$

Furthermore, we assume that conditioned on $\mathbf{z}_{i,j}$, the pixel $\mathbf{y}_{i,j}$ has a multivariate Gaussian density, i.e., for $k = 1, 2, \ldots, K$,

$$p(\mathbf{y}_{i,j} \mid \mathbf{z}_{i,j} = \mathbf{e}_k, \Phi) = \frac{e^{-\frac{1}{2}(\mathbf{y}_{i,j} - \mathbf{m}_k)^T \mathbf{C}_k^{-1} (\mathbf{y}_{i,j} - \mathbf{m}_k)}}{(2\pi)^{3/2} |\mathbf{C}_k|^{1/2}}, \tag{9.3}$$

where $\mathbf{e}_k$ is a K-dimensional binary indicator vector with the kth component being 1. From this, $\Phi = \{\mathbf{m}_k, \mathbf{C}_k\}_{k=1}^K$ contained the mean vectors and covariance matrices for the K image classes. For $\mathbf{z}$, we have adopted an MRF model with a Gibbs' distribution [149]:

$$p(\mathbf{z} \mid \Theta) = \frac{1}{Z} e^{-\beta E(z)}, \tag{9.4}$$

where

$$E(\mathbf{z}) = \frac{1}{2} \sum_{i,j} \sum_{(k,l) \in \mathbf{N}_{i,j}} \left(1 - 2\mathbf{z}_{i,j}^t \mathbf{z}_{k,l}\right) \tag{9.5}$$

is the energy function which decreased (causing $p(\mathbf{z} \mid \Theta)$ to increase) when neighboring pixels were classified into the same class (the set of the neighbors of (i, j) is denoted as $\mathbf{N}_{i,j}$).

Since Θ was generally not sensitive to particular images, it was set manually here. Φ, on the other hand, was directly dependent on the input image and hence had to be estimated during the segmentation process. In this work, as it in [149], this was achieved by using the EM algorithm [177], which amounts to iterating between the following two steps:

1. *E step:* Compute

$$Q(\Phi \mid \hat{\Phi}^{(p)}) = \langle \log p(\mathbf{y} \mid \mathbf{z}, \Phi) + \log p(\mathbf{z} \mid \Theta) \mid \mathbf{y}, \hat{\Phi}^{(p)} \rangle.$$

2. *M step:* Update parameter estimate

$$\Phi^{(p+1)} = \arg\max_{\Phi} Q(\Phi \mid \Phi^{(p)}).$$

Here $\langle \cdot \rangle$ represents the expectation, or mean, and the superscript p denotes the pth iteration. This translated into the following formulas for updating the

parameter estimates:

$$\langle \mathbf{z}_{i,j}^{(p)} \rangle = \sum_{\mathbf{z}_{i,j}} \mathbf{z}_{i,j} f(\mathbf{z}_{i,j}) \tag{9.6}$$

$$\hat{\mathbf{m}}_k^{(p+1)} = \frac{\sum_{i,j} \langle z_{i,j_k}^{(p)} \rangle \mathbf{y}_{i,j}}{\sum_{i,j} \langle z_{i,j_k}^{(p)} \rangle},$$

$$\hat{\mathbf{C}}_k^{(p+1)} = \frac{\sum_{i,j} \langle z_{i,j_k}^{(p)} \rangle [\mathbf{y}_{i,j} - \hat{\mathbf{m}}_k^{(p+1)}][\mathbf{y}_{i,j} - \hat{\mathbf{m}}_k^{(p+1)}]^T}{\sum_{i,j} \langle z_{i,j_k}^{(p)} \rangle}, \tag{9.7}$$

where $k = 1, 2, \ldots, K$, z_{i,j_k} is the kth component of $\mathbf{z}_{i,j}$, and $f(\mathbf{z}_{i,j})$ a "mean field" probability distribution (see Zhang [149]).

These formulas, in addition to providing the estimate of Φ, also produced a segmentation. Specifically, at each iteration, $\langle z_{i,j_k}^{(p)} \rangle$ was interpreted as the probability that $\mathbf{y}_{i,j}$ was assigned to class k. Hence, after a sufficient number of iterations, we can obtain the segmentation $\mathbf{z}$ for each $(i, j) \in \mathbf{L}$ by

$$\mathbf{z}_{i,j} = \mathbf{e}_{k_0} \quad \text{if } k_0 = \arg\max_{1 \le k \le K} \langle z_{i,j_k} \rangle. \tag{9.8}$$

In this way, the EM procedure described above generated the segmentation as a by-product, and therefore provided an alternative to the MAP solution of Eq. (9.1) For those interested in the application of MRF in medical imaging, see Kapur [150], who recently developed a brain segmentation technique which was an extension to the EM work of Wells *et al.* [174] by adding the Gibbs' model to the spatial structure of the tissues in conjunction with a mean-field (MF) solution technique, called Markov random field (MRF) technique (for details on MRF, see Li [151] and Geman and Geman [179]). Thus the technique was named expectation maximization–mean field (EM–MF) technique. By Gibbs' modeling of the homogeneity of the tissue, resistance to thermal noise in the images was obtained. The image data and intensity correction were coupled by an external field to an Ising-like tissue model, and the MF equations were used to obtain posterior estimates of tissue probabilities. This method was more general to the EM-based method and is computationally simple and an inexpensive relaxation-like update. For other work in the area of MRF for brain segmentation see Held *et al.* [152].

Pros and Cons of MRF with Scale Space. The major *advantages* of MRF-based classification is (1) *addition of Gibbs' Model:* Kapur *et al.* model the *a priori* assumptions about the hidden variables as a Gibbs' random field. Thus the prior probability is modeled using the following physics-based analogous Gibbs'

equation: $P(f) = \frac{1}{Z}\exp(-\frac{1}{\kappa T})E(f))$, where $Z = \Sigma_{f'}\exp(\frac{-E(f')}{T})$ and $P(f)$ is the probability of the configuration f, T is the temperature of the system, κ is the Boltzmann constant, and Z is the normalizing constant. The major *disadvantages* of MRF-based classification include the following: (**1**) The computation time would be large if the number of classes is large. In such cases, one needs to use the multiresolution techniques to speed up the computation. (**2**) The positions of the initial clusters are critical in solving the MRF model for the convergence. Here, the initial cluster was computed using K-means algorithm which was a good starting point. However, a more robust method is desirable.

9.3.1.1 Implementation of MRF System

Figure 9.10 shows the main MRF classification system implementation. Input is a perturbed gray scale image with left and right lumens. The number of classes, the initial mean, and the error threshold are inputs given to the system. The result is a classified image with multiple class regions in it, including multiple classes in the lumen region. Figure 9.11 shows the MRF system in more detail. Given the initial center, the number of classes K, the Markov parameters of mean, variance, and covariance, and the perturbed image, the current cluster center is calculated. Using the EM algorithm, new parameters are solved and new cluster

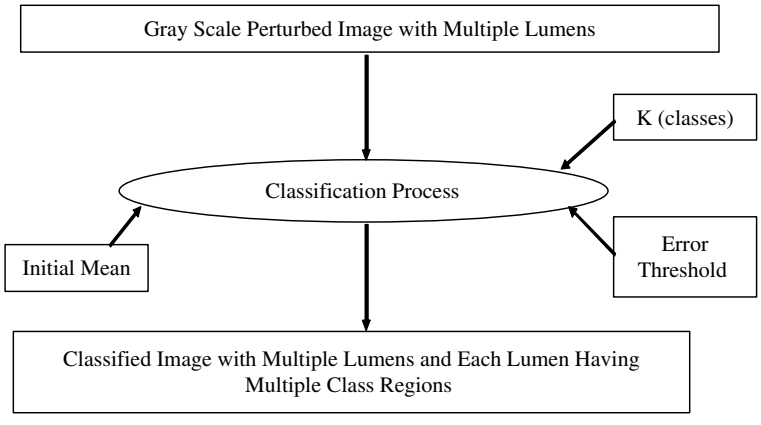


Figure 9.10: Markov random fields (MRF) classification process. Input is a gray scale perturbed image with left and right lumens. The number of classes, the initial mean, and the error threshold are inputs given to the system. The result is a classified image with multiple class regions in it.

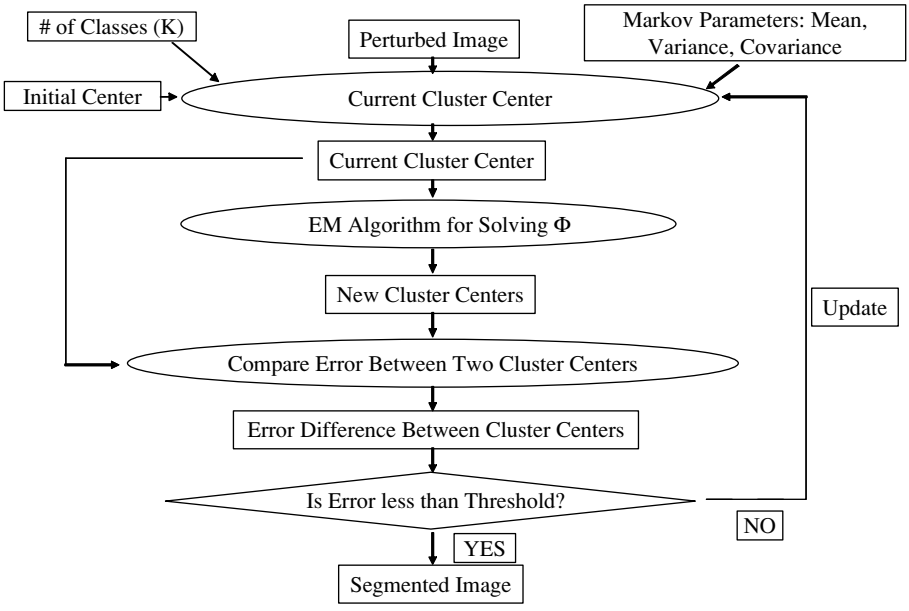


Figure 9.11: The Markov random fields (MRF) segmentation system. Given the initial center, the number of classes K, the Markov parameters of mean, variance, and covariance, and the perturbed image, the current cluster center is calculated. Using the expectation-maximization (EM) algorithm, new parameters are solved and new cluster centers are computed. The error between the previous cluster center and the recently calculated cluster center are compared, and the process is repeated if the error is not less than the error threshold. After the iterative process is finished, the output is a segmented image.

centers are computed. The error between the previous cluster center and the recently calculated cluster center are compared, and the process is repeated if the error is not less than the error threshold. After the iterative process is finished, the output is a segmented image.

9.3.2 Fuzzy-Based Segmentation Method

In this step, we classified each pixel. Usually, the classification algorithm expects one to know how many classes (roughly) the image would have. The number of classes in the image would be the same as the number of tissue types. A pixel could belong to more than one class, and therefore we used the fuzzy membership function to associate with each pixel in the image. There are several

algorithms for computing membership functions, and one of the most efficient ones is Fuzzy C means (FCM) based on the clustering technique. Because of its ease of implementation for spectral data, it is preferred over other pixel classification techniques. Mathematically, we expressed the FCM algorithm below but for complete details, readers are advised to see Bezdek and Hall [180] and Hall and Bensaid [181]. The FCM algorithm computed the measure of membership termed as the *fuzzy membership function*. Suppose the observed pixel intensities in a multispectral image at a pixel location j is given as

$$\mathbf{y_j} = [y_{j1} y_{j2}, \ldots, y_{jN}]^T, \tag{9.9}$$

where j takes the pixel location, and N is the total number of pixels in the data set[4] in FCM (see Figs. 9.12 and 9.13) the algorithm iterates between computing the *fuzzy membership function* and the centroid of each class. This membership function is the pixel location for each class (tissue type), and the value of the membership function lies between the range of 0 and 1. This membership function actually represents the degree of similarity between the pixel vector at a pixel location and the centroid of the class (tissue type); for example, if the membership function has a value close to 1, then the pixel at the pixel location is close to the centroid of the pixel vector for that particular class. The algorithm can be presented in the following four steps. If $u_{jk}^{(p)}$ is the membership value at location j for class k at iteration p, then $\Sigma_{k=1}^{3} u_{jk} = 1$. As defined before, $\mathbf{y}_j$ is the observed pixel vector at location j and $\mathbf{v}_k^{(p)}$ is the centroid of class k at iteration p. Thus, the FCM steps for computing the fuzzy membership values are as follows:

1. Choose the number of classes (K) and the error threshold ϵ_{th}, and set the initial guess for the centroids $\mathbf{v}_k^{(0)}$ where the iteration number $p = 0$.

2. Compute the fuzzy membership function, given by the equation

$$u_{jk}^{(p)} = \frac{\|\mathbf{y}_i - \mathbf{v}_k^{(p)}\|^{-2}}{\sum_{l=1}^{K} \|\mathbf{y}_i - \mathbf{v}^{(p)}\|^{-2}} \tag{9.10}$$

   where $j = 1, \ldots, M$ and $k = 1, \ldots, K$.

3. Compute the new centroids, using the equation

$$\mathbf{v}^{(p+1)} = \frac{\sum_{j=1}^{N} \left(u_{jk}^{(p)}\right)^2 \mathbf{y}_i}{\sum_{j=1}^{N} \left(u_{jk}^{(p)}\right)^2}. \tag{9.11}$$

[4]This is not the $\mathcal{N}$ used in derivation in section 9.4.1.

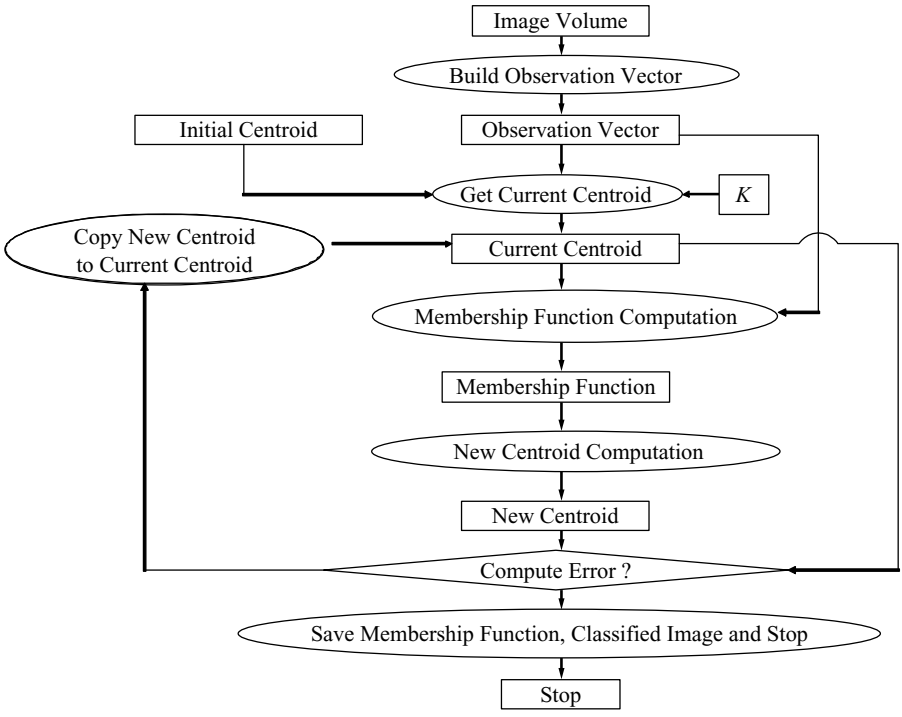


Figure 9.12: Fuzzy C mean (FCM) algorithm. Input is an image volume. An observation vector is built. Initially, the current centroid is given by the initial input centroid and K the number of classes. With the observation vector the membership function is computed, and with it a new centroid is computed. This new centroid is compared to the current centroid, and if the error is too large, the new centroid is copied into the current centroid and the process repeats. Otherwise, if the error is below the threshold, the membership function is saved, and the result is a classified image.

4. Convergence was checked by computing the error between the previous and current centroids ($\|\mathbf{v}^{(p+1)} - \mathbf{v}^{(p)}\|$). If the algorithm had converged, an exit would be required; otherwise, one would increment p and go to step 2 for computing the fuzzy membership function again. The output of the FCM algorithm was K sets of fuzzy membership functions. We were interested in the membership value at each pixel for each class. Thus, if there were K classes, then we threw out K number of images and K number of matrices for the membership functions to be used in computing the final speed terms.

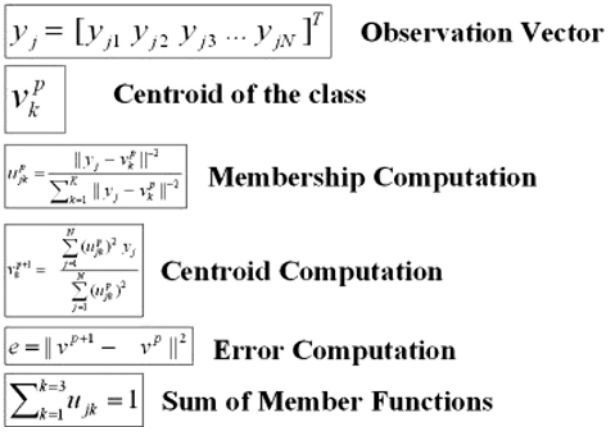


Figure 9.13: Mathemetical expression of the FCM algorithm. Equations for the observation vector, centroid of the class, sum of the membership function, membership computation, centroid computation, and error computation are shown.

9.3.3 Graph-Based Segmentation Method

The graph segmentation method (GSM) segments an image by treating it as a graph $G = (V, E)$ where V the set of vertices are the pixels and E the set of edges are pairs of pixels. Using a weight function $w(e)$, where e is an edge $(v_i,\ v_j)$, the weights of the edges are computed and the edges are sorted by weight in a nondecreasing order. Initially, each pixel v_i is segmented into its own component C_i.

For each edge $(v_i,\ v_j)$ in the list, a decision criterion D is applied and a decision is made whether or not to merge the components C_i and C_j. After this decision is made on each edge in the list, the result is a list of the final components of the segmented image.

The input image is first smoothed by a given smoothing parameter σ. Input constant k determines the size preference of the components by changing the *threshold function* $\tau(C)$ (see Figs. 9.14 and 9.15).

The decision criteria D is a comparison between the *difference* between components C_i and C_j and the *minimum internal difference* among C_i and C_j. The difference between two components is defined as the minimum weight of the edges that connect the two components:

$$Dif(C_i, C_j) = \min_{v_i \in C_i, v_j \in C_j, (v_i, v_j) \in E} w((v_i, v_j)). \tag{9.12}$$

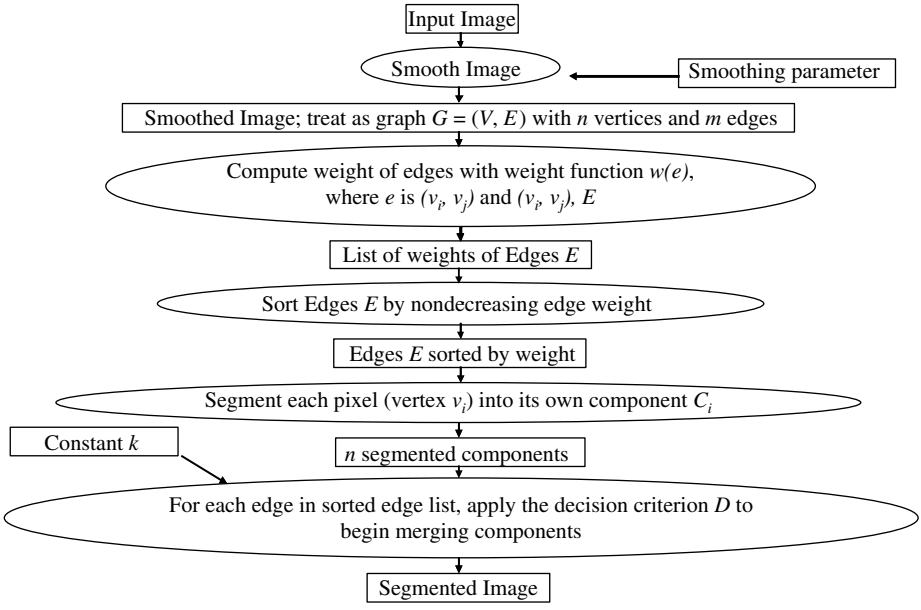


Figure 9.14: Graph segmentation method (GSM). The input image is smoothed given a smoothing parameter. The image is treated as a graph, with each pixel treated like a vertex. An edge is a pair of pixels. Using a weight function $w(e)$, the weights of the edges are computed and the edges are listed by weight in a nondecreasing order. Initially, each pixel is segmented into its own component. For each edge in the list, a decision criterion D is applied and the components are merged accordingly. Input constant k determines the size preference of the components by changing the threshold function. The result is a segmented image made up of the final merged components.

The minimum internal difference among two components C_i and C_j is defined as the minimum of the sum of the *internal difference* and the threshold function of each component:

$$MInt(C_i, C_j) = \min(Int(C_i) + \tau(C_i), Int(C_j) + \tau(C_j)), \tag{9.13}$$

where the internal difference $Int(C)$ of a component C is defined as the maximum weight in the minimum spanning tree $MST(C, E)$ of the component:

$$Int(C) = \max_{e \in MST(C,E)} w(e), \tag{9.14}$$

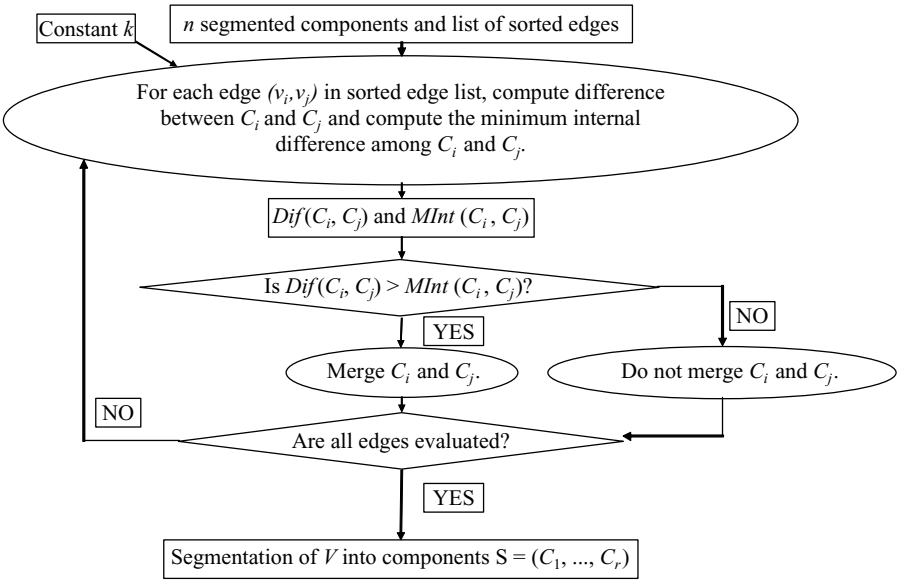


Figure 9.15: Decision criterion D for the graph segmentation method (GSM). After the list of edge weights are sorted and each pixel is segmented into its own component, the decision criterion is D is applied to each edge. The constant k is used in determining the threshold function. First the difference between the two components to which the two pixels making up the edge belong is computed. Then the minimum internal difference among those two components is computed. If the difference between the two components is greater than the minimum internal difference among them, then D applied to the two components is true, and the two components are not merged because there is evidence for a boundary between them. Otherwise, if the difference between the two components is less than or equal to the minimum internal difference, then the D applied to the two components is false, and the two components are merged into one component which contains both pixels of the edge. This decision criterion is applied to all the edges of the list, and the final result is a segmentation of the pixels into components.

and where the threshold function $\tau(C)$ is defined as

$$\tau(C) = \frac{k}{|C|}, \tag{9.15}$$

where k is the input constant and $|C|$ is the size of the component C.

$$Int(C) = \max_{e \in MST(C,E)} w(e)$$ Internal Difference

$$Dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w((v_i, v_j))$$ Difference between two components

$$\tau(C) = \frac{k}{|C|}$$ Threshold Function

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2))$$ Minimum Internal Difference

Figure 9.16: Graph segmentation method (GSM) equations. The internal difference of a component is the maximum edge weight of the edges in its minimum spanning tree. The difference between two components is the minimum edge weight of the edges formed by two pixels, one belonging to each component. The threshold function of a component is the constant k divided by the size of that component, where the size of a component is the number of pixels it contains. The minimum internal difference among two components is the minimum value of the sum of the internal difference and the value of the threshold function of each component.

If the difference between the two components is greater than the minimum internal difference among the two components, then the two components are not merged. Otherwise, the two components are merged into one component.

9.4 Synthetic System Design and Its Processing

9.4.1 Image Generation Process

The model equation for generation of the observed image is shown in Eq. (9.16).

$$\mathbf{I}_{\text{observe}} = \mathbf{I}_{\text{original}} + \eta \tag{9.16}$$

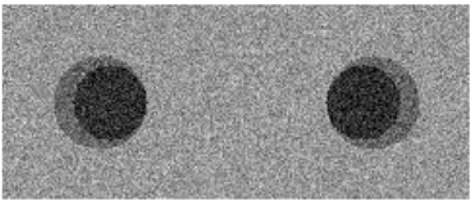


Figure 9.17: Synthetic pipeline with $\sigma^2 = 500$.

This can be expressed for every pixel as

$$I_{\text{observe}}(x, y) = I_{\text{original}}(x, y) + \eta(x, y) \tag{9.17}$$

where $\eta(x, y) \sim \mathcal{N}(0, \sigma^2)$ and σ^2 is the variance of the noise. $\mathcal{N}$ is the Gaussian distribution. The output synthetic image using Gaussian image generation process is shown in the Fig 9.17.

Figure 9.18 shows eight different directions that the core class of the lumen can be with respect to the crescent moon class. The darkest region is the core

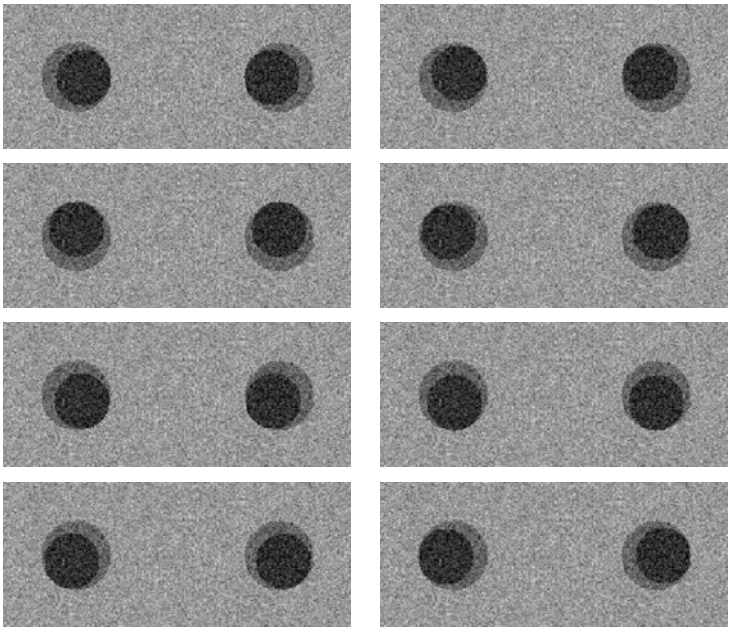


Figure 9.18: $\sigma^2 = 500$, all directions, large noise protocol. With respect to the center of the lumen area, the core class is shown in eight different orientations. In the top row, from right to left: east, northeast, north; in the second row: northwest, southeast, south; in the third row: southwest, west, west.

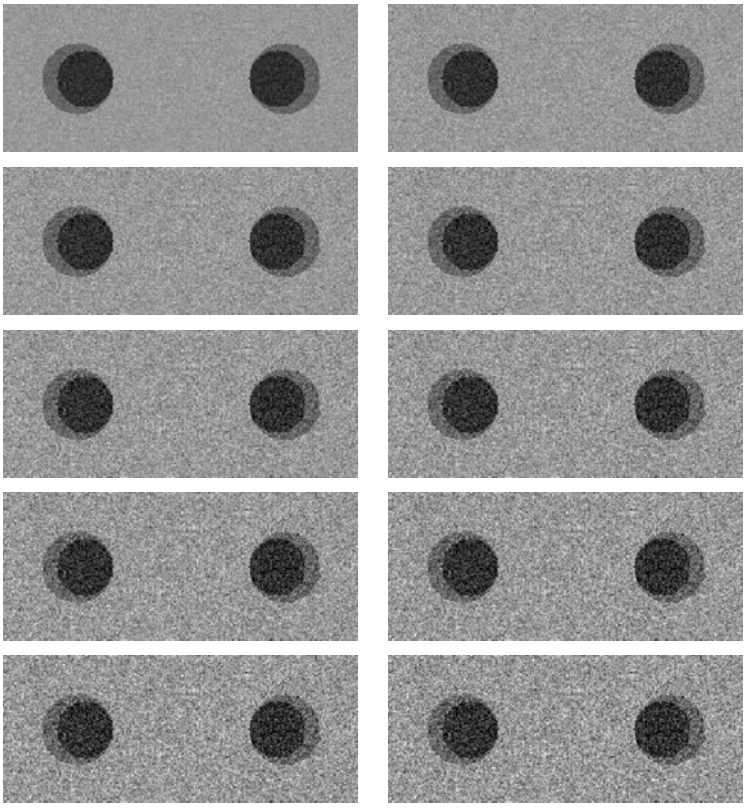


Figure 9.19: Images with 10 different variances using large noise protocol. The gray scale model is perturbed with variance (σ^2) varying from 100 to 1000. In the top row, from right to left: $\sigma^2 = 100$ and 200; in the second row: $\sigma^2 = 300$ and 400; in the third row: $\sigma^2 = 500$ and 600; in the fourth row: $\sigma^2 = 700$ and 800; in the fifth row: $\sigma^2 = 900$ and 1000.

lumen, and the next lightest region that surrounds the core is the crescent moon class. Figure 9.19 shows the core class and the crescent moon class of the lumen with perturbation. The darkest region is the core lumen, and the next lightest region that surrounds the core is the crescent moon class. The variance (σ^2) was varied from 100 to 1000.

9.4.2 Lumen Detection and Quantification System

System pipeline is shown in Fig. 9.20. Step one consists of synthetic generation process discussed in section 9.4.1. This consists of synthesizing the two lumens corresponding to the left and right side of the neck. The grayscale image

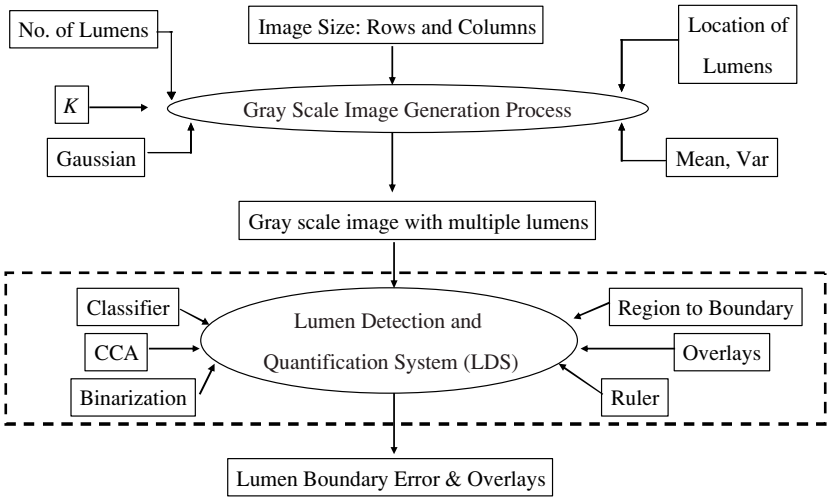


Figure 9.20: Block diagram of the system. A gray scale image is generated, with parameters being the number of lumens, location of lumens, the number of classes K, and a Gaussian perturbation with mean and variance. The result is an image with multiple lumens with noise. This image is then processed by the lumen detection and quantification system (LDS). This system includes many steps including classification, binarization, connected components analysis (CCA), boundary detection, overlaying, and measuring the error. The final result are the lumen errors and overlays.

generation process takes in the noise parameters: the mean and variance, the locations of the lumens, the number of lumens and the class intensities of the lumen core, the crescent moon, and the background.

Step two consists of lumen detection and quantification system (LDAS) (see Fig. 9.20). The major block is the classification system discussed in section 9.3. Then comes the binarization unit which is used to convert the classified input into the binarized image and also does the region merging. It also has a connected component analysis (CCA) system block which is the input to the LDAS. We also need the region-to-boundary estimator which will give the boundary of the left and right lumens. Finally we have the quantification system (called Ruler), which is used to measure the boundary error.

The LDQS system consists of lumen detection and lumen quantification system. The lumen detection system (LDS) is shown in Fig. 9.21. The detection process is done by the classification system, while the identification is done by the CCA system. There are three classification system we have used in our

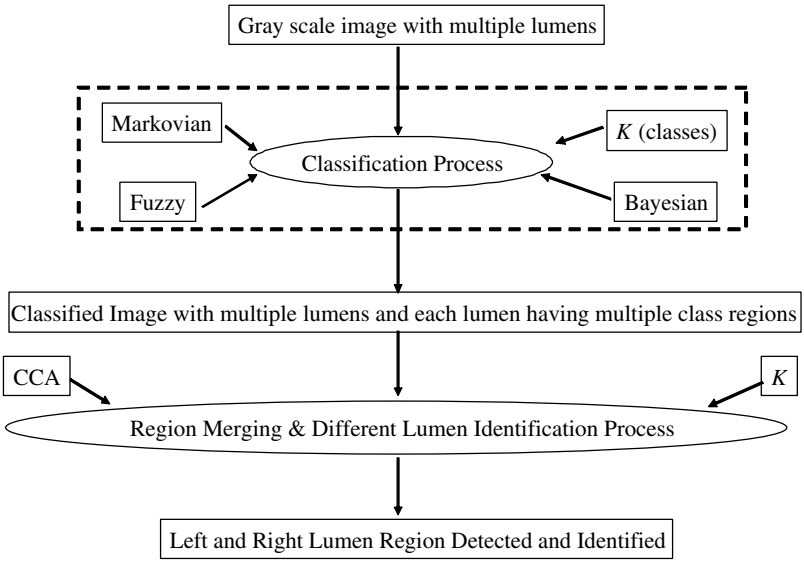


Figure 9.21: Block diagram of lumen detection system (LDS). The gray scale image with multiple lumens is first classified by one of the classifiers. The result is a classified image with multiple lumens, with each lumen having multiple class regions. Within each lumen, these multiple regions are merged in the binarization process, given the number of classes K. They are labeled using connected component analysis (CCA). The LDS detects and labels each lumen.

processes (see section 9.3). The parameters used are the number of classes (K) as shown in the Fig. 9.21. The CCA block also takes the parameter the number of classes, K, as input.

The lumen detection and identification is further detailed as shown in the Fig. 9.22. The detection system inputs the classified image and outputs the binary regions of the lumen. Because of boundary classes and plaque diffusion in the lumen area, there are classes well. We merge these classes to generate the complete lumen and the final detection of the lumen takes place as shown in the Fig. 9.22. Finally, the system shows the identification of the left and right lumen using the CCA analysis.

9.4.3 Region Merging for Lumen Detection

Figure 9.23 shows how the regions with multiple classes are merged. We will discuss the region merging strategy a little differently for the real data analysis,

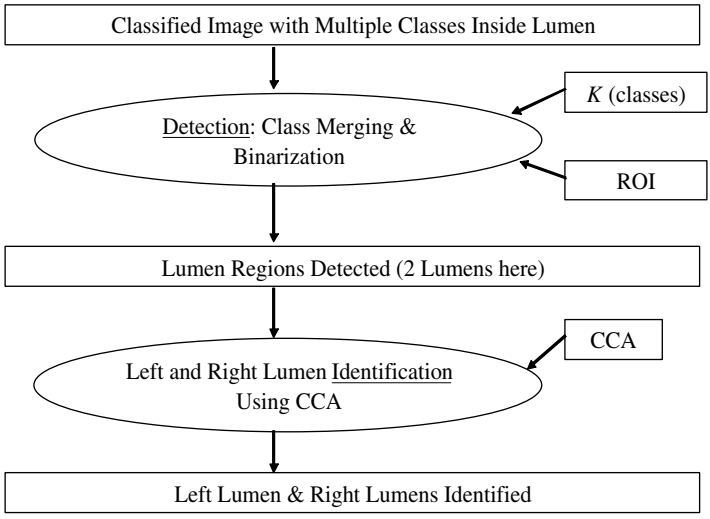


Figure 9.22: Detection and identification of lumen. Input image is a classified image with multiple classes inside the lumens. Given the number of classes K and the region of interest (ROI) of each region, the appropriate classes are merged and the image is binarized. The detected lumens are then identified using connected component analysis (CCA), and the left lumen and right lumen are identified.

due to the bifurcations in the arteries of the plaqued vessels (see sections 9.6.1 and 9.6.2). Figure 9.23 illustrates the region merging algorithm. The input image has lumens which have one, two, or more classes. If the number of classes in the ROI is one class, then that class is selected; if two classes are in the ROI, then the minimum class is selected; and if there are three or more classes in the ROI, then the minimum two classes are selected. The selected classes are merged by assigning all the pixels of the selected classes one level value. This process results in the binarization of the left and right lumens.

The binary region labeling process is shown in Fig. 9.24. The process uses the CCA approach of top to bottom and left to right. Input is an image in which the lumen regions are binarized. The CCA first labels the image from the top to the bottom, and then from the left to the right. The result is an image that is labeled from the left to the right.

ID assignment process of the CCA for each pixel is shown in Fig. 9.25. In the CCA, in the input binary image, each white pixel is assigned a unique ID. The label propagation process then results in connected components. The propagation of

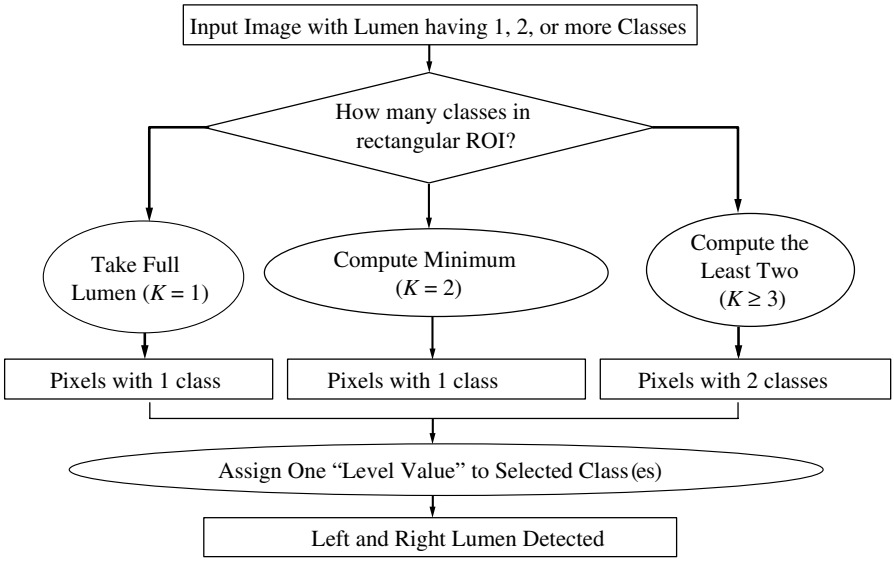


Figure 9.23: Region detection: region merging algorithm. The input image has lumens that have 1, 2, or more classes. If the number of classes in the ROI is one class, then that class is selected; if two classes are in the ROI, then the minimum class is selected; and if there are three or more classes in the ROI, then the minimum two classes are selected. The selected classes are merged by assigning all the pixels of the selected classes one level value. This process results in the left and right lumen being binarized.

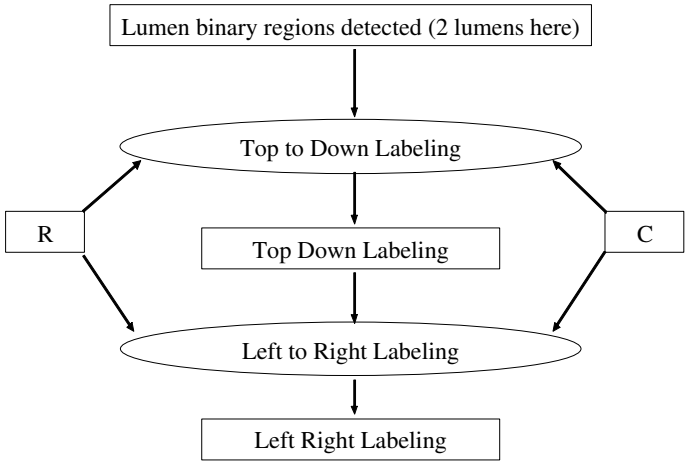


Figure 9.24: Region identification using connected component analysis (CCA). Input is an image in which the lumen binary regions are detected. The CCA first labels the image from the top to the bottom, and then from the left to the right. The result is an image that is labeled from the left to the right.

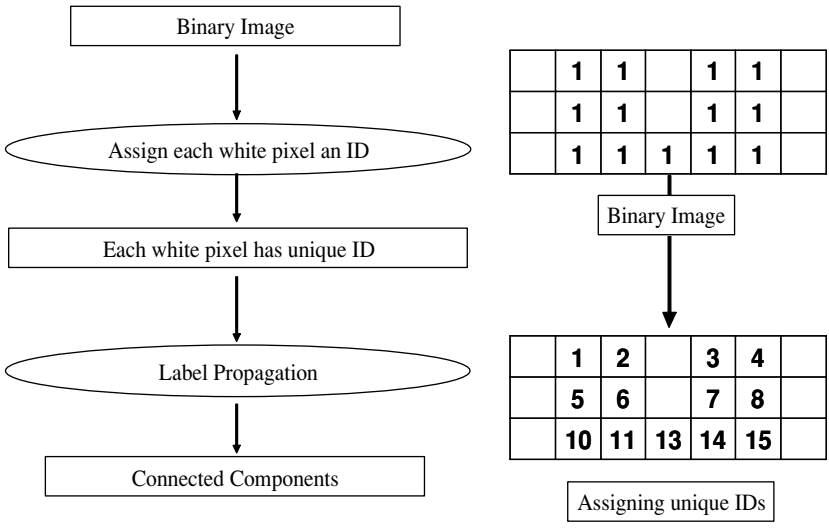


Figure 9.25: Region identification: ID assignment. In the connected component analysis (CCA), in the input binary image each white pixel is assigned a unique ID. Then the label propagation process results in connected components.

the region from left to right is shown in Fig. 9.26. This is the first pass of the label-propagation process. Every row of the image is scanned from top to bottom, left to right, pixel by pixel. If the pixel has an ID, then pixels to the left and above of the pixel are checked for IDs, and if either one has an ID, then the pixel's value is reassigned to that of the lowest among the neigbor pixels and the pixel. This processed is repeated for all pixels in the row and in all rows. The result is a binary image with some label propagation. The propagation of the region from top to bottom is shown in Fig. 9.27. This is the second pass of the label-propagation process. Every row of the image is scanned from bottom to top, right to left, pixel by pixel. If the pixel has an ID, then pixels to the right and below of the pixel are checked for IDs, and if either one has an ID, then the pixel's value is reassigned to that of the lowest among the neigbor pixels and the pixel. This processed is repeated for all pixels in the row and in all rows. The result is a binary image with some label propagation. Finally, the region assignment is summarized in Fig. 9.28. The top left image is a binary image with a value of 1 assigned to each of the white pixels. Each of the white pixels are assigned a unique value in the top-right image. The left to right and top to bottom label propagation propagates the labels of value 1 and 3, and the result is the

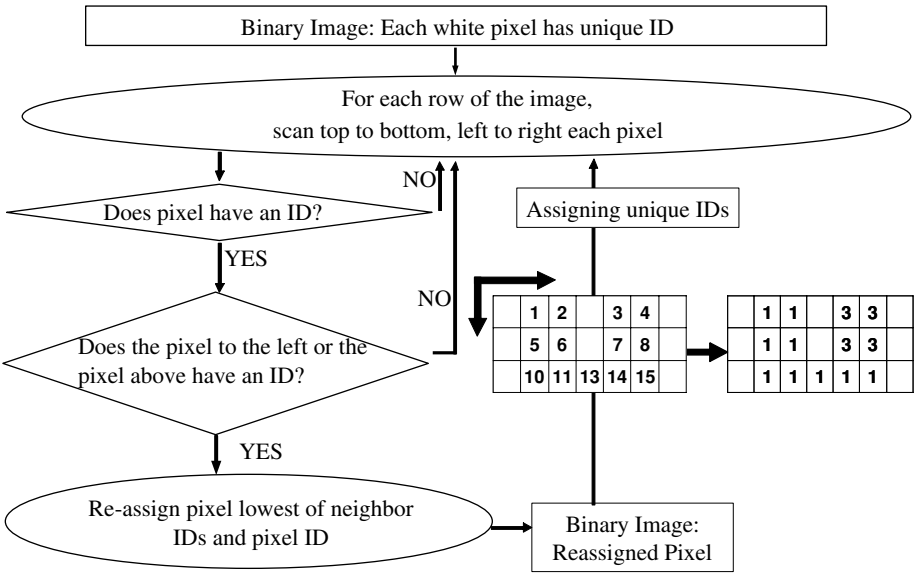


Figure 9.26: Region identification: propagation. This is the first pass of the label propagation process. Given the bianary image having unique IDs for each white pixel, every row of the image is scanned from top to bottom, left to right, pixel by pixel. If the pixel has an ID, then pixels to the left and above of the pixel are checked for IDs, and if either one has an ID, then the pixel's value is reassigned to that of the lowest among the neigbor pixels and the pixel. This processed is repeated for all pixels in the row all rows. The result is a binary image with reassigned pixel values.

bottom-left image. Then, the right to left and bottom to top label propagation propagates the label value of 1 to the pixels having a label value of 3. The result is the bottom-right image, in which the connected white pixels have all the same label values of 1. This is the basic algorithm of the process; the CCA we used uses look-up tables in order to efficiently assign regions in two passes. The results on CCA on a binary image with 4 lumens are shown in Fig. 9.29. The input image has the lumens detected, but they are all of the same color. CCA identifies the lumens by labeling each with a different color. The process to generate a color image is shown in Fig. 9.30. The first input is a gray scale image. The second input is the ideal boundary image. This image is dilated and converted to a red color, resulting in a red ideal boundary image. The third input is the estimated boundary image. This image is dilated and converted to a green color, resulting

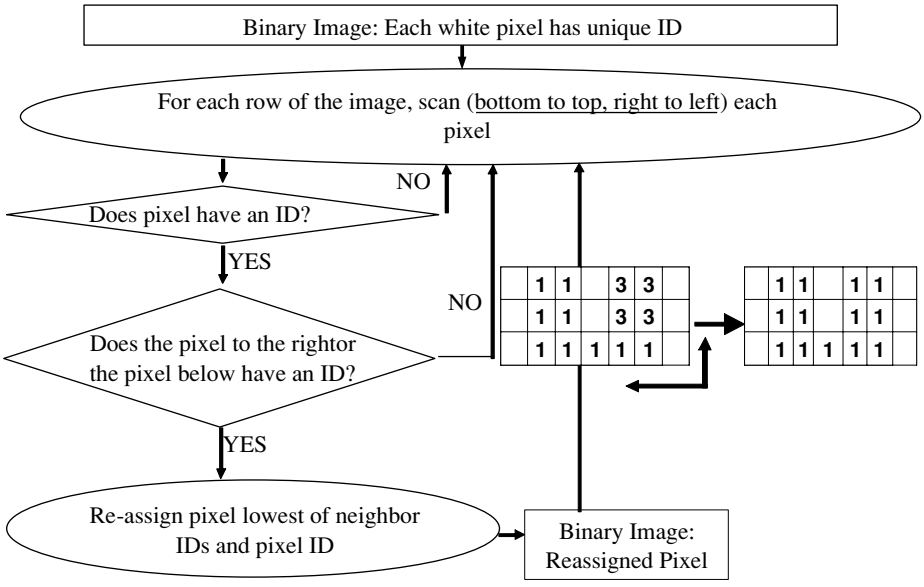


Figure 9.27: Region identification: propagation. This is the second pass of the label propagation process. Given the binary image having unique IDs for each white pixel, every row of the image is scanned from bottom to top, right to left, pixel by pixel. If the pixel has an ID, then pixels to the right and below of the pixel are checked for IDs, and if either one has an ID, then the pixel's value is reassigned to that of the lowest among the neigbor pixels and the pixel. This processed is repeated for all pixels in the row and all rows. The result is a binary image with reassigned pixel values.

in green estimated boundary image. These three images are fused to produce a color overlay image.

9.4.4 Results of Synthetic System: Boundary Estimation

Figure 9.31 shows in the FCM classification system all the steps for the left and right lumen detection, identification, and boundary estimation process in the synthetic images. We look at large noise protocol as an example below with noise level $\sigma^2 = 500$. In the first row the left image shows the synthetically generated image. In the first row the right image shows the image after it has been smoothed by the Perona–Malik smoothing function. In the second row the left image shows the classified image after the image has gone through the FCM

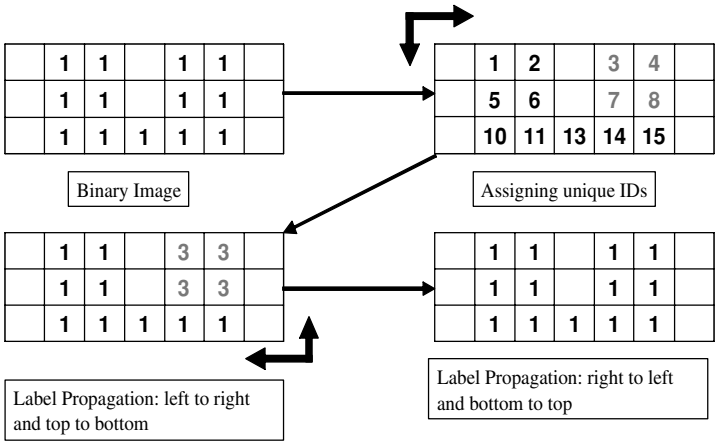


Figure 9.28: Region identification: ID Propagation. The top left image is a binary image with a value of 1 assigned to each of the white pixels. Each of the white pixels are assigned a unique value in the top right image. The left to right and top to bottom label propagation propagates the labels of value 1 and 3, and the result is the bottom left image. Then, the right to left and bottom to top label propagation propagates the label value of 1 to the pixels having a label value of 3. The result is the bottom right image, in which the connected white pixels have all the same label values of 1.

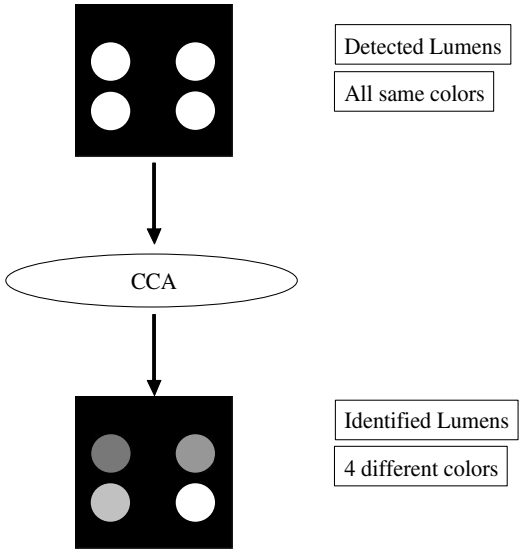


Figure 9.29: Region identification: CCA. The input image has the lumens detected, but they are all the same color. Connected component analysis (CCA) identifies the lumens by labeling each with a different color.

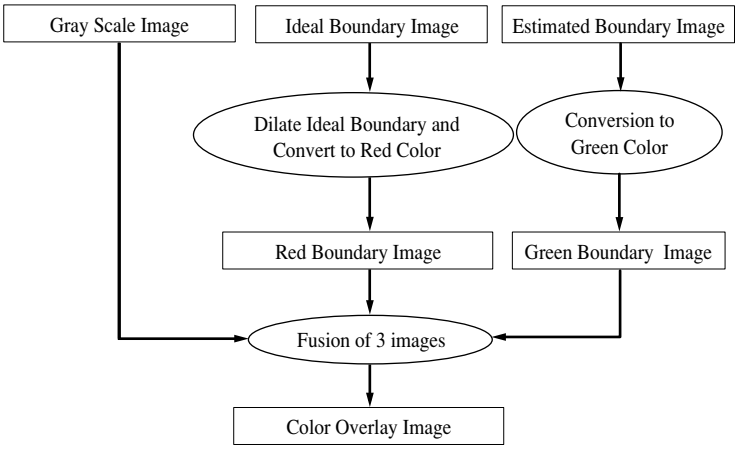


Figure 9.30: Color overlay block. The first input is a gray scale image. The second input is the ideal boundary image. This image is dilated and converted to a red color, resulting in a red ideal boundary image. The third input is the estimated boundary image. This image is dilated and converted to a green color, resulting in green estimated boundary image. These three images are fused to produce a color overlay image.

classification system. In the second row the right image shows the binarization of the image after selecting only the core class for binarization ($K = 1$). In the third row the left image shows the binarization of the image after selecting the core class and the edge classes for binarization ($K > 1$). In the third row the right image shows the image ($K = 1$) after the labeling of CCA. In the fourth row the left image shows the image ($K > 1$) after the labeling of CCA. In the fourth row the right image shows the image ($K = 1$) after the labeling of assign ID. In the fifth row the left image shows the image ($K > 1$) after the labeling of assign ID. In the fifth row the right image shows the computer-estimated boundary of the image ($K = 1$) using the region-to-boundary algorithm. In the sixth row the left image shows the computer-estimated boundary of the image ($K > 1$) using the region-to-boundary algorithm. In the sixth row the right image shows the original image on which is overlayed the ideal ground truth boundary, the artifacted boundary ($K = 1$), and the corrected boundary ($K > 1$).

Figure 9.32 shows in the MRF classification system all the steps for the left and right lumen detection, identification, and boundary estimation process in the synthetic images. We look at large noise protocol as an example below with

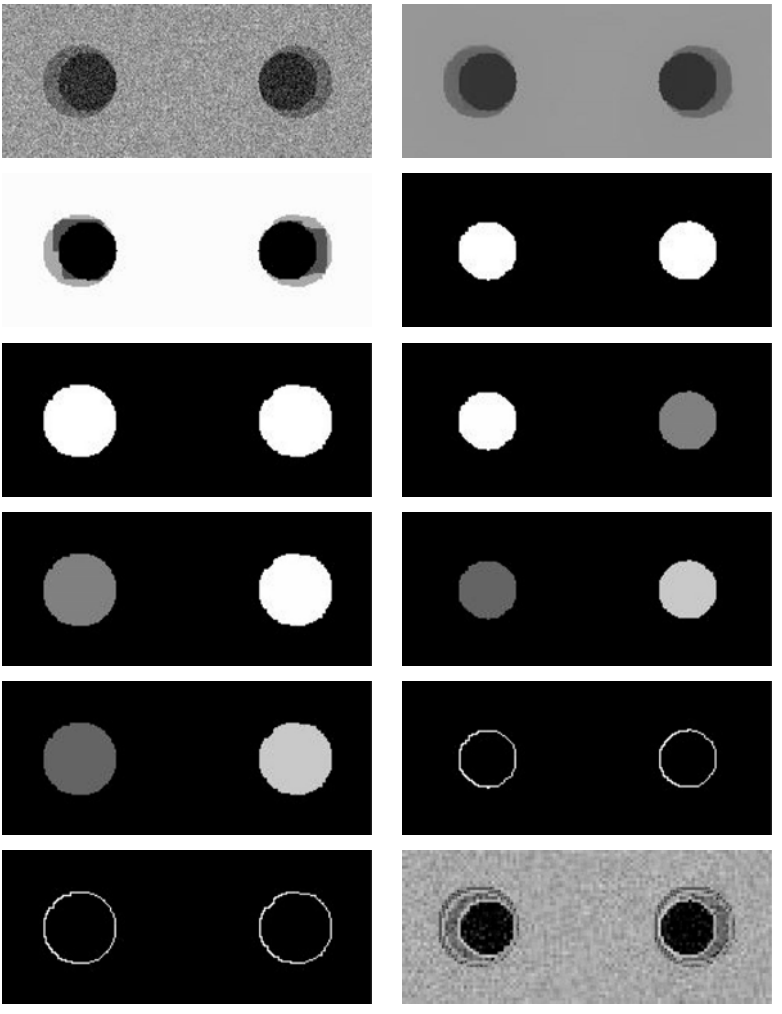


Figure 9.31: Results on synthetic image with noise variance, $\sigma^2 = 500$ using FCM method. Row 1, left: Synthetic generate image. Row 1, right: After Perona–Malik smoothing. Row 2, left: After FCM classification system. Row 2, right: Binarization with only **C0** class ($K = 1$). Row 3, left: Binarization with merging **C0**, **C1**, and **C2** classes ($K > 1$). Row 3, right: Binarization after CCA ($K = 1$). Row 4, left: Binarization after CCA ($K > 1$). Row 4, right: After assign ID ($K = 1$). Row 5, left: After assign ID ($K > 1$). Row 5, right: After region to boundary ($K = 1$). Row 6, left: After region to boundary ($K > 1$). Row 6, right: Overlay generation with and without crescent moon.

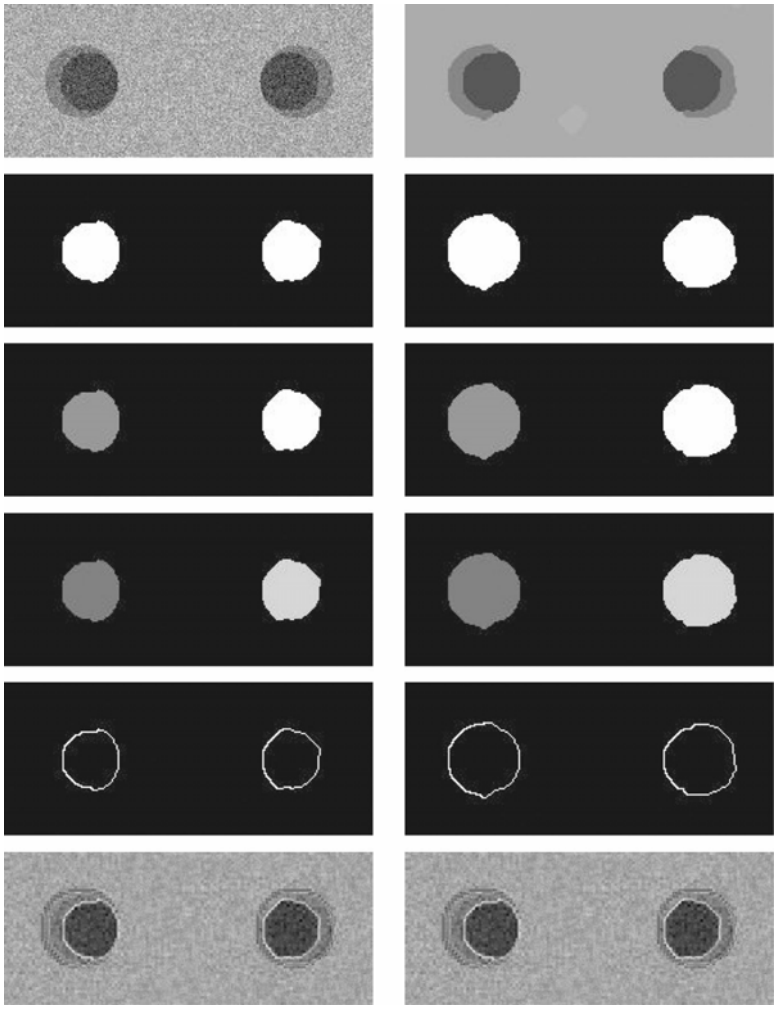


Figure 9.32: Results on synthetic image with noise variance, $\sigma^2 = 500$ using MRF method. Row 1, left: Synthetic generate image. Row 1, right: After MRF classification system. Row 2, left: Binarization with only **C0** class ($K = 1$). Row 2, right: Binarization with merging **C0**, **C1**, and **C2** classes ($K > 1$). Row 3, left: Binarization after CCA ($K = 1$). Row 3, right: Binarization after CCA ($K > 1$). Row 4, left: After assign ID ($K = 1$). Row 4, right: After assign ID ($K > 1$). Row 5, left: After region to boundary ($K = 1$). Row 5, right: After region to boundary ($K > 1$). Row 6, left: Overlay generation with and without crescent moon. Row 6, right: Overlay generation with and without crescent moon.

noise level $\sigma^2 = 500$. In the first row the left image shows the synthetically generated image. In the first row the right image shows the classified image after the image has gone through the MRF classification system. In the second row the left image shows the binarization of the image after selecting only the core class for binarization ($K = 1$). In the second row the right image shows the binarization of the image after selecting the core class and the edge classes for binarization ($K > 1$). In the third row the left image shows the image ($K = 1$) after the labeling of CCA. In the third row the right image shows the image ($K > 1$) after the labeling of CCA. In the fourth row the left image shows the image ($K = 1$) after the labeling of assign ID. In the fourth row the right image shows the image ($K > 1$) after the labeling of assign ID. In the fifth row the left image shows the computer-estimated boundary of the image ($K = 1$), using the region-to-boundary algorithm. In the fifth row the right image shows the computer-estimated boundary of the image ($K > 1$), using the region-to-boundary algorithm. In the sixth row the left image shows the original image on which is overlayed the ideal ground truth boundary, the artifacted boundary ($K = 1$), and the corrected boundary ($K > 1$). In the sixth row the right image shows the original image on which is overlayed the ideal ground truth boundary, the artifacted boundary ($K = 1$), and the corrected boundary ($K > 1$).

Figures 9.33 and 9.34 show in the GSM classification system all the steps for the left and right lumen detection, identification, and boundary estimation process in the synthetic images. We look at large noise protocol as an example below with noise level $\sigma^2 = 500$. In Fig. 9.33, the first row the left image shows the synthetically generated image. In the first row the right image shows the image after it has been smoothed by the Perona–Malik smoothing function. In the second row the left image shows the image after its frequency peaks of pixel values have been merged. In the second row the right image shows the classified image after the image has gone through the GSM classification system. In the third row the left image shows the binarization of the image after selecting only the core class for binarization ($K = 1$). In the third row the right image shows the binarization of the image after selecting the core class and the edge classes for binarization ($K > 1$). In the fourth row the left image shows the image ($K = 1$) after the labeling of CCA. In the fourth row the right image shows the image ($K > 1$) after the labeling of CCA. In Fig. 9.34, the first row the left image shows the image ($K = 1$) after the labeling of assign ID. In the first row the right image shows the image ($K > 1$) after the labeling of assign

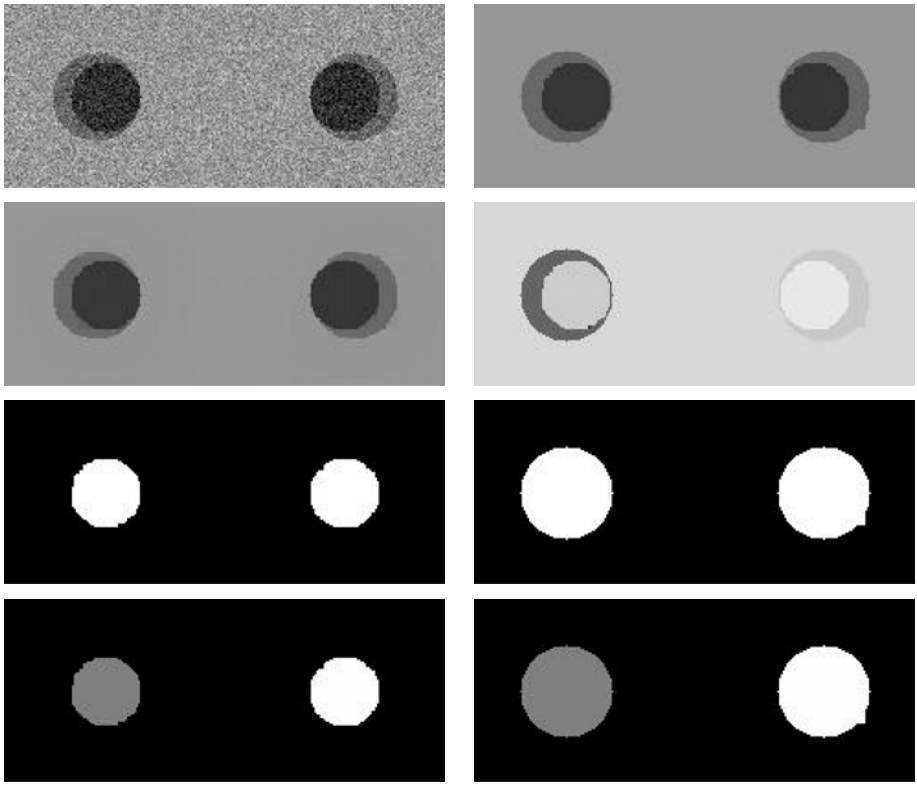


Figure 9.33: Results on synthetic image with noise variance, $\sigma^2 = 500$ using GSM method. Row 1, left: Synthetic generate image. Row 1, right: After peak merger. Row 2, left: After Perona–Malik Smoothing. Row 2, right: After GSM classification system. Row 3, left: Binarization with only **C0** class ($K = 1$). Row 3, right: Binarization with merging **C0**, **C1**, and **C2** classes ($K > 1$). Row 4, left: Binarization after CCA ($K = 1$). Row 4, right: Binarization after CCA ($K > 1$).

ID. In the second row the left image shows the computer-estimated boundary of the image ($K = 1$), using the region-to-boundary algorithm. In the second row the right image shows the computer-estimated boundary of the image ($K > 1$), using the region-to-boundary algorithm. In the third row the left image shows the original image on which is overlayed the ideal ground truth boundary, the artifacted boundary ($K = 1$), and the corrected boundary ($K > 1$). In the third row the right image shows the original image on which is overlayed the ideal ground truth boundary, the artifacted boundary ($K = 1$), and the corrected boundary ($K > 1$).

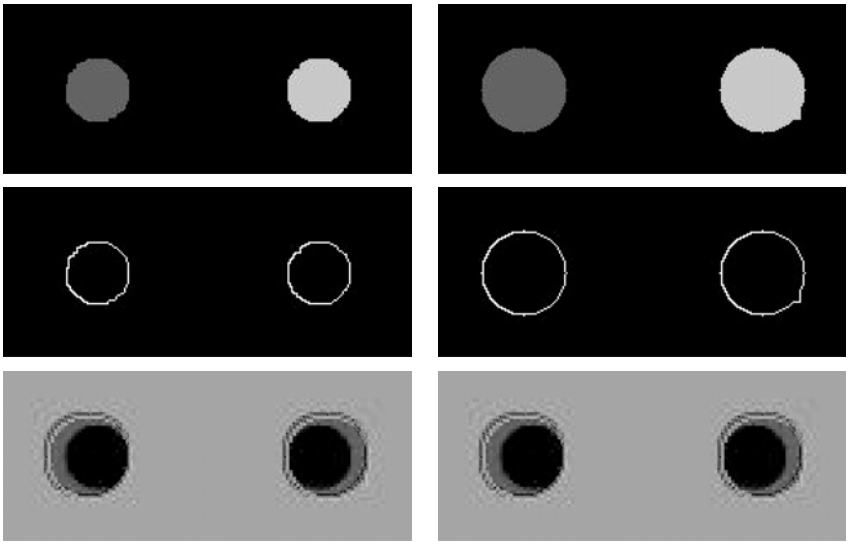


Figure 9.34: Results on synthetic image with noise variance, $\sigma^2 = 500$ using GSM method. Row 5, left: After assign ID ($K = 1$). Row 5, right: After assign ID ($K > 1$). Row 6, left: After region to boundary ($K = 1$). Row 6, right: After region to boundary ($K > 1$). Row 7, left: Overaly generation with and without crescent moon. Row 8, right: Overaly generation with and wihout crescent moon.

9.5 Performance Evaluation System: Rulers and Error Curves

The polyline distance $D_s(B_1 : B_2)$ between two polygons representing boundary B_1 and B_2 is symmetrically defined as the average distance between a vertex of one polygon and the boundary of the other polygon. To define this measure precisely, we first need to define a distance $d(v, s)$ between a point v and a line segment s. The distance $d(v, s)$ between a point v having coordinates (x_0, y_0), and a line segment having end points (x_1, y_1) and (x_2, y_2) is

$$d(v,s) = \begin{cases} \min\{d_1, d_2\} & \text{if } \lambda < 0, \lambda > 1 \\ |d^{\perp}| & \text{if } \leq 0, \lambda \leq 1, \end{cases} \tag{9.18}$$

where

$$d_1 = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}$$
$$d_2 = \sqrt{(x_0 - x_2)^2 + (y_0 - y_2)^2}$$

$$\lambda = \frac{(y_2 - y_1)(y_0 - y_1) + (x_2 - x_1)(x_0 - x_1)}{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$
$$d^{\perp} = \frac{(y_2 - y_1)(x_1 - x_0) + (x_2 - x_1)(y_0 - y_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \tag{9.19}$$

The distance $d_b(v, B_2)$ measuring the polyline distance from vertex v to the boundary B_2 is defined by

$$d_b(v, B_2) = \min_{s \in sides B_2} d(v, s) \tag{9.20}$$

The distance $d_{vb}(B_1, B_2)$ between the vertices of polygon B_1 and the sides of polygon B_2 is defined as the sum of the distances from the vertices of the polygon B_1 to the closest side of B_2.

$$d_{vb}(B_1, B_2) = \sum_{v \in vertices\, B_1} d(v, B_2)$$

Reversing the computation from B_2 to B_1, we can similarly compute $d_{vb}(B_2, B_1)$. Using Eq. (9.20), the polyline distance between polygons, $D_s(B_1 : B_2)$ is defined by

$$D_s(B_1 : B_2) = \frac{d_{vb}(B_1, B_2) + d_{vb}(B_2, B_1)}{(\#\text{ vertices} \in B_1 + \#\text{ vertices} \in B_2)} \tag{9.21}$$

9.5.1 Mean Error (e_{NFP}^{poly})

Using the definition of the polyline distance between two polygons, we can now compute the mean error of the overall system. It is denoted by e_{NFP}^{poly} and defined by

$$e_{NFP}^{\text{poly}} = \frac{2 \times \sum_{t=1}^{F} \sum_{n=1}^{N} D_s(G_{nt}, C_{nt})}{F \times N} \tag{9.22}$$

where $D_s(G_{nt}, C_{nt})$ is the polyline distance between the ground truth G_{nt} and computer-estimated polygons C_{nt} for patient study n and slice number t. Using the definition of the polyline distance between two polygons, the standard deviation can be computed as

$$\sigma_{NFP}^{\text{poly}} = \sqrt{\frac{\sum_{t=1}^{F} \sum_{n=1}^{N} \{\sum_{v \in \text{vertices}\, G_{nt}} (d_b(v, C_{nt}) - e_{NFP}^{\text{poly}})^2 + \sum_{v \in \text{vertices}\, C_{nt}} (d_b(v, G_{nt}) - e_{NFP}^{\text{poly}})^2\}}{N \times F \times (\#\text{ vertices} \in B_1 + \#\text{ vertices} \in B_2)}} \tag{9.23}$$

9.5.2 Error per Vertex and Error per Arc Length for Bias Computation

Using the polyline distance formulaes, we can compute the error per vertex from one polygon (ground truth) to another polygon (computer estimated). This is defined as the mean error for a vertex v over all the patients and all the slices. The *error per vertex* for a fixed vertex v when computed between ground truth and computer-estimated boundary is defined by

$$e_v^{GC} = \frac{\sum_{t=1}^{F} \sum_{n=1}^{N} d_b(v, G_{nt})}{F \times N} \tag{9.24}$$

Similary we can compute the *error per vertex* between computer estimated and ground truth using Eq.(9. 20) *Error per arc length* is computed in the following way: For the values e_v^{GC} where $v = 1, 2, 3, \ldots, P_1$, we construct a curve f^{GC} defined on the interval [0,1] which takes the value e_v^{GC} at point x which is the *normalized arc length* to vertex v and whose in between values are defined by linear interpolation. We compute the curve f^{CG} between computer estimated boundary and ground truth boundary in a similar way. We then add algebraically these two curves to yield the final *error per arc length*, given as $f = \frac{f^{GC}+f^{CG}}{2}$.

9.5.3 Performance of Synthetic System: Error Curves

9.5.3.1 Small Noise Protocols

Figure 9.35 (left and right) shows the performance of the synthetic system for the small noise protocol using polyline (see section 9.5) and shortest distance methods. Figure 9.35 (left) compares the mean error curves of the MRF vs. FCM (with smoother) using the PDM, while Fig. 9.35 (right) compares the mean error curves of the MRF vs. FCM (with smoother) using the shortest distance method. Using PDM, as the variance of the noise (σ^2) increases from 0 to 100, the mean error in both methods increases gradually. The mean error for the FCM (with smoother) remains under 1.6 pixels, while the mean error for MRF ranges between 1.6 and 1.8 pixels. The same pattern is observed using SDM method (see Fig. 9.35, right). It is also seen in the two graphs that FCM using PDM has a lower error compared to FCM using SDM.

In another procotol, we run the same PDM and SDM for FCM methods but with and without the Perona–Malik smoothing process. This can be seen in Fig. 9.36 (left and right). The method of PDE-based smoothing system improves

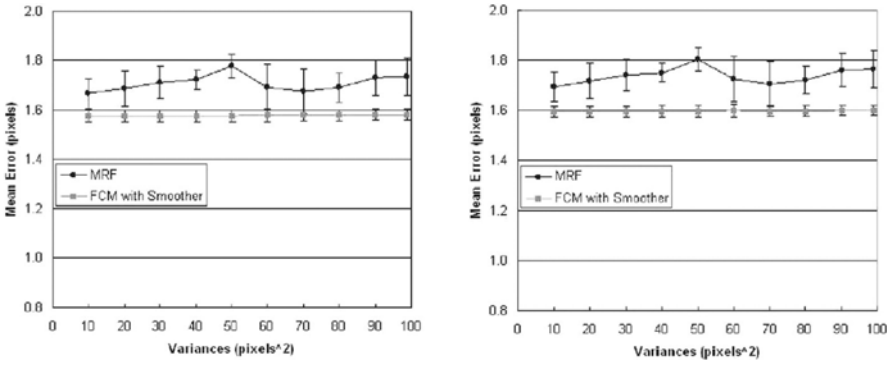


Figure 9.35: Results of MRF vs. FCM using PDM and SDM methods for small noise protocol. Left: MRF vs. FCM using PDM method. Right: MRF vs. FCM using SDM method.

the error over non-PDE based system at large noise and thus is more robust in identification and detection process. It is also seen in the two graphs that FCM (with and without smoother) using PDM has a lower error compared to FCM (with and without smoother) using SDM.

In another procotol, we compare the MRF vs. FCM (without PDE smoother) and this can be seen in Fig. 9.37 (left and right). Using PDM, as the variance of the noise (σ^2) increases from 0 to 100, the mean error in both methods increases gradually. The mean error for the FCM (without smoother) remains under 1.6 pixels, while the mean error for MRF ranges between 1.6 and 1.8 pixels. The same pattern is observed using SDM method (see Fig. 9.37, right). It is also seen

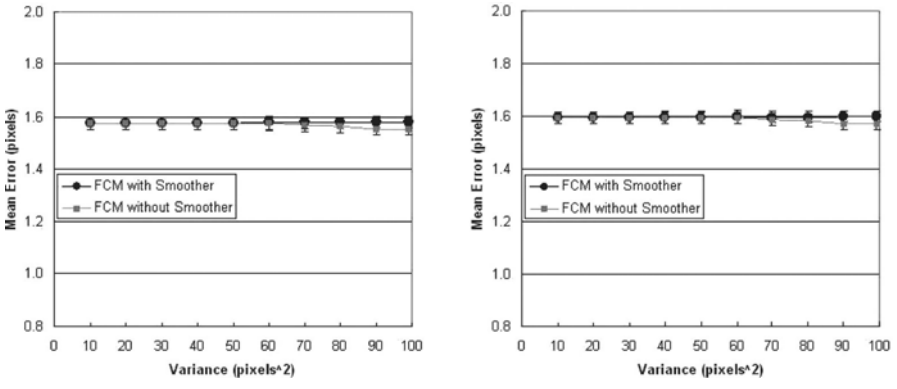


Figure 9.36: Effect of PDE-smoother process on the overall system. Left: FCM using PDM method (for small noise protocol). Right: FCM using SDM method (for small noise protocol).

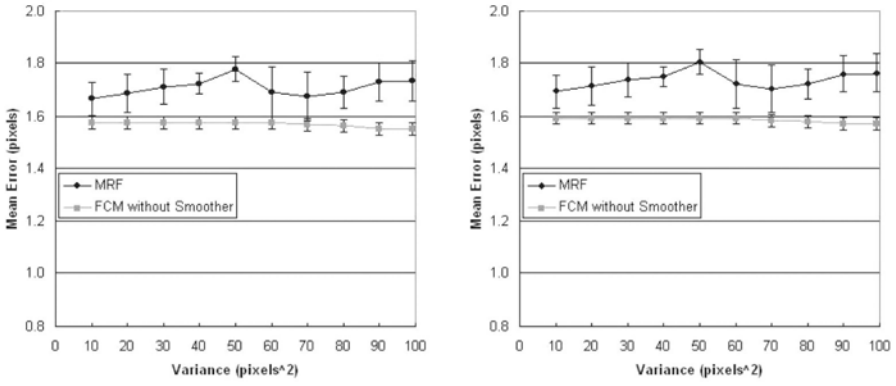


Figure 9.37: MRF vs. FCM. Left: MRF using PDM method (for small noise protocol). Right: FCM using PDM method (for small noise protocol).

in the two graphs that FCM (without smoother) and MRF using PDM have a lower error compared to FCM (without smoother) and MRF using SDM.

9.5.4 Large Noise Protocols

In one of the protocols, we compare MRF vs. FCM (with smoother) using PDM (see Fig. 9.38, left) and SDM (see Fig. 9.38, right). Using PDM, as the variance of the noise (σ^2) increases from 100 to 1000, the mean error in FCM increases very gradually. This clearly demonstrates its robustness at large noise variances. The

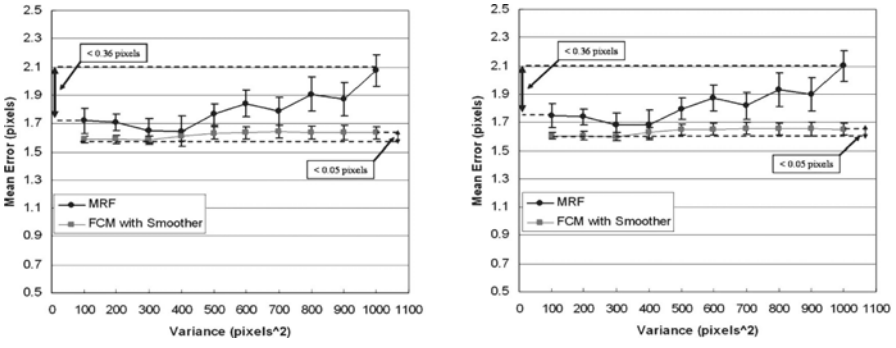


Figure 9.38: MRF vs. FCM for large noise protocol. Left: MRF using PDM method. Right: FCM using PDM method. Note that the range of the mean errors is less than 0.36 pixels using MRF and less than 0.05 pixels using FCM with smoother.

mean error for MRF increases more rapidly compared to that of FCM. The mean error for the FCM (with smoother) remains close to 1.6 pixels, while the mean error for MRF ranges between 1.7 ($\sigma^2 = 100$) and 2.1 pixels ($\sigma^2 = 1000$). The same pattern is observed for the MRF vs. FCM using SDM (see Fig. 9.38, right).

9.5.5 Bias Estimation Protocol

From section 9.5.2, we compute the error per vertex (point) around the boundary consisting of 150 points for MRF and FCM (with smoother) methods. We used the PDM ruler for bias error analysis. It can be seen from Fig. 9.39 that there is no bias error for the FCM method, and all the boundary points have an error with a mean of 1.5 pixels. For the MRF method, the bias error curve first becomes negative, and then rises to positive values after the boundary point 45. This shows that there has been a right shift of the computer-estimated contour compared to the ideal contour after the point 45. This also means that the one third of the contour is inside the ideal boundary, and the remaining two thirds of the contour is outside the ideal boundary. Such a behavior can be explained using two concepts: (a) intra- and interobserver variability and (b) shift in the estimated contour. This is out of the scope of this chapter and will be discussed elsewhere.

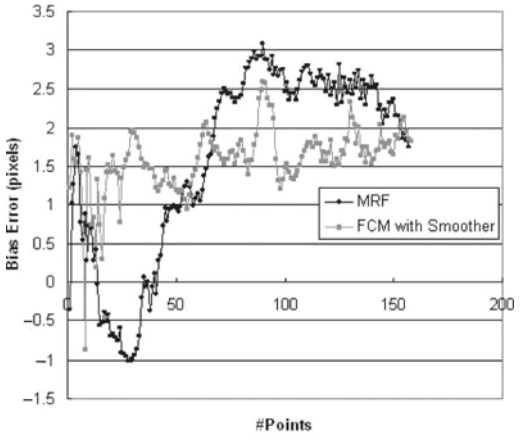


Figure 9.39: The bias error is compared between the MRF and the FCM with smoother. The mean errors are plotted against consecutive points around the contour. Large noise protocol.

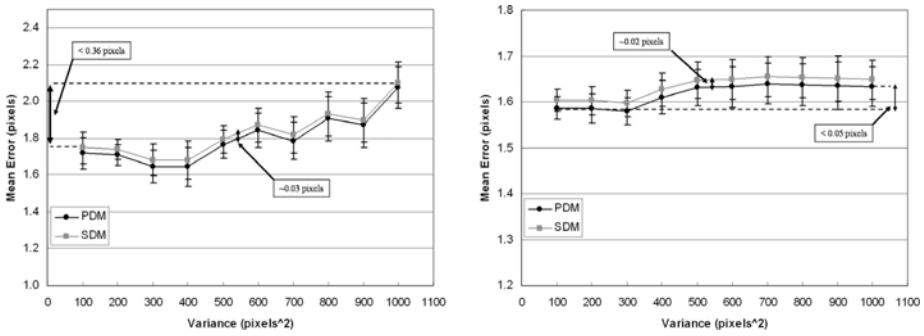


Figure 9.40: PDM vs. SDM methods. Left: MRF: PDM vs. SDM. Right: FCM: PDM vs. SDM. The length of the range of the mean errors is less than 0.36 pixels, and the difference between the two curves is about 0.03 pixels.

9.5.6 PDM versus SDM Performances

Figure 9.40 (left) shows the comparison between PDM and SDM, using MRF classification system for large noise protocol ($\sigma^2 = 100$ to $\sigma^2 = 1000$). As seen in the figure, the mean error increases gradually from 1.7 to 2.1 pixels. As seen in the figure, PDM has a lower error compared to SDM by a small amount of 0.05 pixels. However, they both go hand-in-hand as the variance increases from 100 to 1000.

Figure 9.40 (right) shows the comparison between PDM and SDM, using FCM classification system for large noise protocol (σ^2) from 100 to 1000. As seen in the figure, the mean error increases gradually from less than 1.6 pixels to little more than 1.6 pixels. As seen in the v figure, PDM has a lower error compared to SDM by a small amount of 0.05 pixels. However, they both also go hand-in-hand as the variance increases from 100 to 1000. We also see that FCM method (see Fig. 9.40, right) has a lower mean error slope compared to MRF (see Fig. 9.40, left). The explanation of this can be attributed to the discussion of the previous section.

9.5.7 Shape Optimization Protocol

In this protocol, we study and analyze the shape characteristics of the lumen with respect to the number of points on the boundary. We know that as the number of

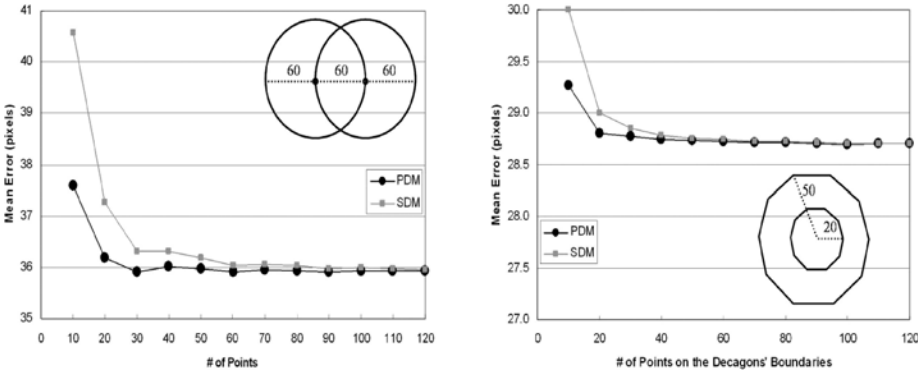


Figure 9.41: Sampling protocol test: Left: Shape optimization test. Right: Concentric shape decagon test. Two circle contours each of radius 60 pixels have their centers separated by 60 pixels. The mean errors given by the PDM and the SDM are plotted against the number of points on the circular contours. As the number of points on each of the contours increases, the difference in the mean errors decreases, and both errors approach an actual value.

points increases on the boundary, the boundary becomes more smooth, but we do not know as to how many points are necessary on the boundary to represent the best lumen shape. Figure 9.41 (left and right) demonstrates the mean error around the boundary versus the number of points on the lumen boundary. As the number of points increases from 10 to 120, the mean error drops rapidly using PDM and SDM methods. Using PDM, the mean error drops rapidly when the number of boundary points increases from 10 to 30 and reaches a stage of convergence when the number of points is 50. The same pattern is observed using the SDM method and the mean error falls rapidly from points 10 to 50 and reaches a stage of convergence when the number of points on the boundary is 80. The stage of convergence here means that there is no more change in the mean error, if the number of points increases beyond a certain limit. Lastly, the fall of the errors as the number of points increases is more rapid for SDM compared to that of PDM, and the starting error (when total points are 10) in SDM is much larger compared to that of PDM. A similar experiment was done synthetically when the boundaries are concentric shapes. We took a simple shape of a concentric decagon (with radius 20 and 50 pixels) and increasing the number of points from 10 to 120. Since the boundaries were concentric, the

point of convergence was same (70 points) for both PDM and SDM (see Fig. 9.41, right).

9.5.8 Protocol for Spline Fitting Over Boundaries

The last protocol consists of fitting a Bezier curve (spline) to the boundaries. There are two-fold purposes in our protocol: (a) To make the boundary curves smoother and (b) to make the points on the boundary curves equidistant. We use the methodology discussed in Graphics Gems for spline-fitting, and we used curve interpolation for making the curve equidistant. Both these effects show a reduction in the mean error. Figure 9.42 shows the effect of

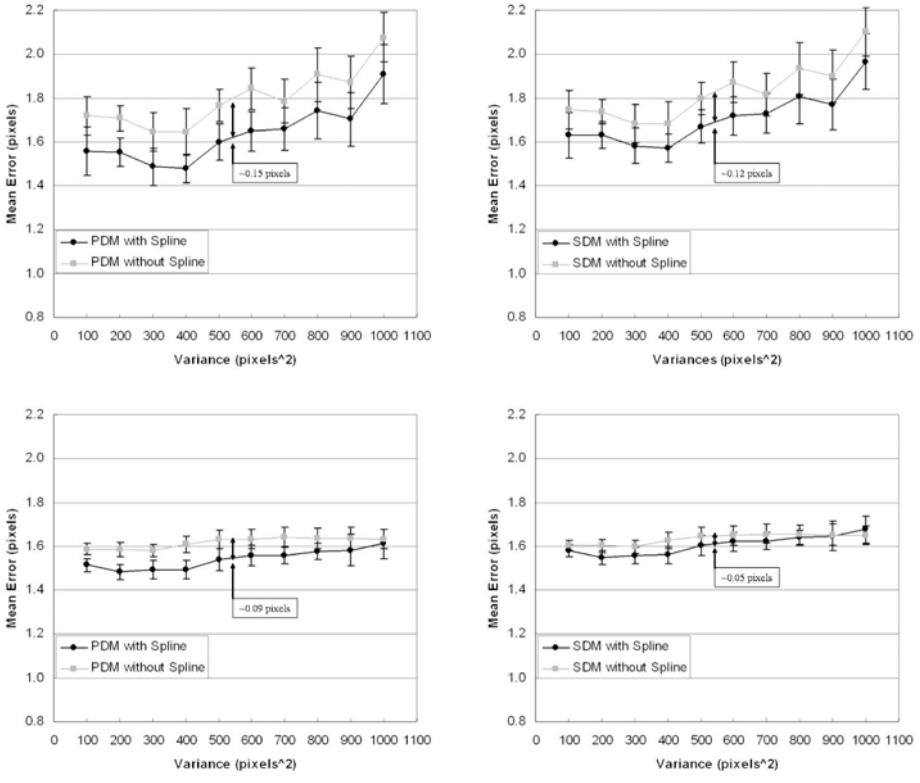


Figure 9.42: Effect of fitting splines over the estimated boundaries. Top left: MRF, PDM, with and without splines. Top right: MRF, SDM, with and without splines. Bottom left: FCM, PDM, with and without splines. Bottom right: FCM, SDM, with and without splines.

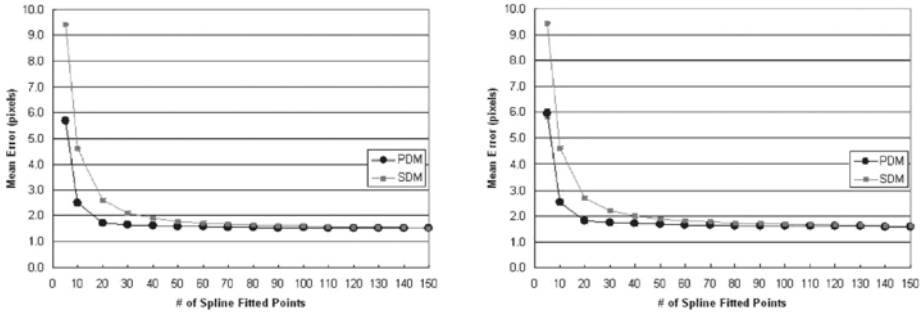


Figure 9.43: Optimization curves. Left: $\sigma^2 = 500$. Right: $\sigma^2 = 1000$.

fitting splines over the estimated boundaries. There are four parts in this figure showing the effect of splines over two classification systems, using two distance methods: (a) MRF using PDM, (b) MRF using SDM, (c) FCM using PDM, and (d) FCM using SDM. In all four subprotocols, we find the same behavior that the spline-fitted mean errors are lower than nonspline-fitted mean errors. We also observed that there is a very consistent standard deviation error for all four subprotocols. We also did lumen shape optimization on fitted spline shapes, and this can be seen in Fig. 9.43 (left and right). As the number of points on the boundary increases, the mean error drops and reaches a stage of convergence.

9.6 Real Data Analysis: Circular Vs. Elliptical

9.6.1 Circular Binarization

Select class is used for binarization of the classified image. The frequency of each pixel value in the ROI is determined. The core class **C0** is the class with the greatest number of pixels. The number of pixels is equivalent to the area of the core. The average area of the entire lumen core is determined from the ground truth boundaries, and this area is compared to the area of the **C0** area. A threshold function is used to determine whether to binarize the **C0** region, or to merge **C0** with **C1** and then binarize. We now discuss the methods to compute the average lumen area, lumen core area, and the difference of these, and their comparison.

9.6.1.1 Lumen Area Computation by Triangle/Scan-Line Methods (A)

To determine the average area of the entire lumen from the ground truth boundaries, the area by triangles computation is used. The center point of the ROI is the user input, and is equivalent to the center of gravity (CG). The area of the enclosed region is obtained by summing the areas of the triangles formed by the CG and each pair of neighboring points on the boundary.

In the scan-line method, we count the number of pixels along the scan line which lies in the ROI. This process is done for all the lies which interest the ROI region. The entry and exit points are computed by finding the number of times the scan line interests the boundary yielding the odd or even number. If the intersection yields 1 then begin counting the pixels, and if the intersection yields 2, then stop counting pixels. This gives a total number of pixels along the line. The process stops when there are no more interesections. In a 384×512 image, the average area for the left and right lumen is 500 pixels squared.

9.6.1.2 Area of Lumen Core Class (B)

The select class package takes as one of its inputs the number of classes formed after the segmentation method. Using this as a size for an array of the different classes **C0** through $\mathbf{C_n}$, the program checks each pixel in the ROI and stores the number of times that each of the different pixel values occur. The program then sorts these class values by their frequency.

9.6.1.3 Difference Computation (A − B) and Comparison with Threshold

Using the average ground truth contour area, a difference threshold, Td, is determined. We set $\mathrm{Td} = 75$. If the difference between the average ground truth contour area and the number of pixels of **C0** in the ROI is less than the difference threshold, then only **C0** is selected. If the difference is greater than the difference threshold, then both **C0** and **C1** are selected, then they are merged and a binary image is made.

In the GSM, a select class package is not used, but a region growing method is used. The GSM usually merges the **C0** and **C1** classes, so the region growing captures both **C0** and **C1** classes.

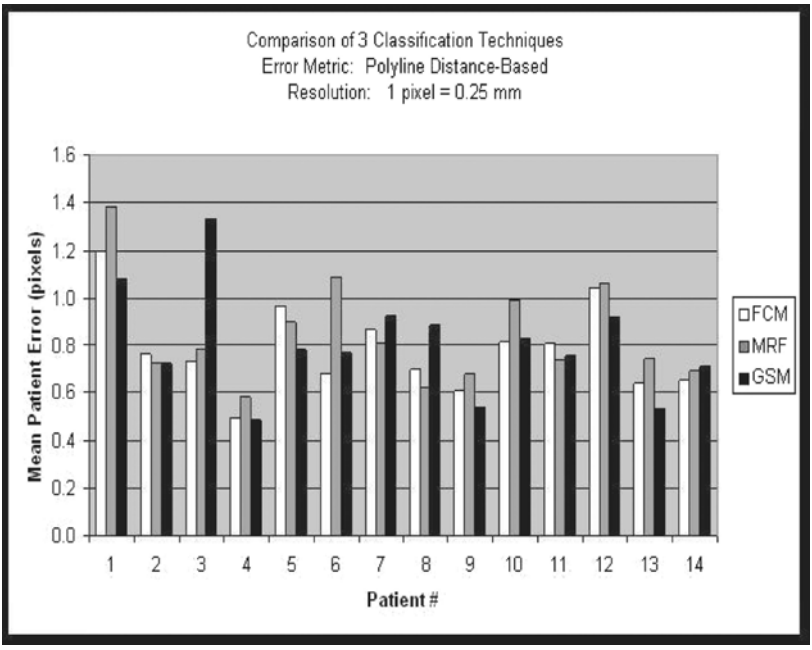


Figure 9.44: Results using FCM, MRF, and GSM methods.

9.6.2 Elliptical Binarization

The equations used for rotation of the ellipse about its center with an angle α is given by the new coordinates in Eq. (9.25).

$$x_i' = (x_i - x_0)\cos(\alpha) - (y_i - y_0)\sin(\alpha) \tag{9.25}$$

$$y_i' = (x_i - x_0)\sin(\alpha) - (y_i - y_0)\cos(\alpha) \tag{9.26}$$

9.6.3 Performance Evaluation of Three Techniques

Figure 9.44 shows the mean error bar charts for the three pipelines (i.e., using three classification systems: MRF, FCM, and GSM methods).[5] The charts can be seen in the Tables 9.1–9.3. Table 9.1 shows the error between the computer-estimated boundary and ground truth boundary using FCM-based

[5] We ran the system using each of the three different classifying methods on real patient data. Ground truth boundaries of the walls of the carotid artery were traced for 15 patients. Overall the number of boundary points was roughly 22,500 points. A pixel was equivalent to 0.25 mm. Using MRF, the average error was 0.61 pixels; using FCM, the average error was 0.62 pixels; using GSM, the average error was 0.74 pixels.

Table 9.1: Mean errors as computed using polyline and shortest distance methods when the classification system is FCM based

Patient No.	Artifacted (PDM)	Corrected (PDM)	Artifacted (SDM)	Corrected (SDM)
1	2.052	1.195	2.063	1.216
2	0.928	0.764	0.948	0.794
3	3.174	0.729	3.180	0.756
4	1.106	0.490	1.118	0.513
5	1.514	0.968	1.529	0.993
6	1.079	0.681	1.094	0.704
7	1.278	0.863	1.310	0.893
8	0.928	0.695	0.944	0.723
9	0.758	0.606	0.783	0.631
10	1.004	0.813	1.027	0.840
11	1.407	0.808	1.418	0.826
12	1.408	1.042	1.426	1.078
13	0.735	0.643	0.753	0.670
14	0.922	0.655	0.939	0.685

method. Column 1 shows the error when the estimated boundary is not corrected (artifacted), using the PDM ruler. Column 2 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen, using the PDM ruler. Column 3 shows the error when the estimated boundary is not corrected (artifacted), using the SDM ruler. Column 4 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen, using the SDM ruler. As seen in the table, column 2 shows the least error and is signifiicantly improved over the artifacted boundaries. Table 9.2 shows the error between the computer-estimated boundary and ground truth boundary using MRF-based method. Column 1 shows the error when the estimated boundary is not corrected (artifacted), using the PDM ruler. Column 2 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen, using the PDM ruler. Column 3 shows the error when the estimated boundary is not corrected (artifacted), using the SDM ruler. Column 4 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen, using the SDM ruler. As seen in the table, column 2 shows the least error and is significiantly improved over the artifacted boundaries. Table 9.3 shows the error between the computer-estimated boundary and ground truth boundary using GSM-based method. Column 1 shows the error when the estimated boundary is not corrected (artifacted), using the PDM ruler. Column 2 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen,

Table 9.2: Mean errors as computed using polyline and shortest distance methods when the classification system is MRF based

Patient No.	Artifacted (PDM)	Corrected (PDM)	Artifacted (SDM)	Corrected (SDM)
1	1.609	1.382	1.627	1.402
2	0.831	0.726	0.857	0.759
3	1.174	0.781	1.195	0.805
4	0.687	0.584	0.706	0.605
5	1.239	0.895	1.263	0.917
6	1.164	1.086	1.182	1.105
7	1.100	0.807	1.124	0.839
8	1.004	0.620	1.023	0.645
9	0.696	0.679	0.714	0.702
10	0.890	0.958	0.912	0.982
11	0.938	0.736	0.954	0.763
12	0.941	1.065	0.965	1.089
13	0.679	0.740	0.704	0.761
14	0.851	0.694	0.869	0.716

Table 9.3: Mean errors as computed using polyline and shortest distance methods when the classification system is GSM based

Patient No.	Artifacted (PDM)	Corrected (PDM)	Artifacted (SDM)	Corrected (SDM)
1	1.081	1.081	1.105	1.105
2	0.721	0.721	0.746	0.746
3	1.329	1.329	1.351	1.351
4	0.487	0.487	0.505	0.505
5	0.778	0.778	0.802	0.802
6	0.767	0.767	0.788	0.788
7	0.920	0.920	0.949	0.949
8	0.885	0.885	0.903	0.903
9	0.536	0.536	0.559	0.559
10	0.826	0.826	0.849	0.849
11	0.752	0.752	0.774	0.774
12	0.914	0.914	0.942	0.942
13	0.533	0.533	0.557	0.557
14	0.708	0.708	0.732	0.732

using the PDM ruler. Column 3 shows the error when the estimated boundary is not corrected (artifacted), using the SDM ruler. Column 4 shows the error when the estimated boundary is corrected by merging multiple classes of the lumen, using the SDM ruler. As seen in the table, column 2 shows the least error and is significiantly improved over the artifacted boundaries.

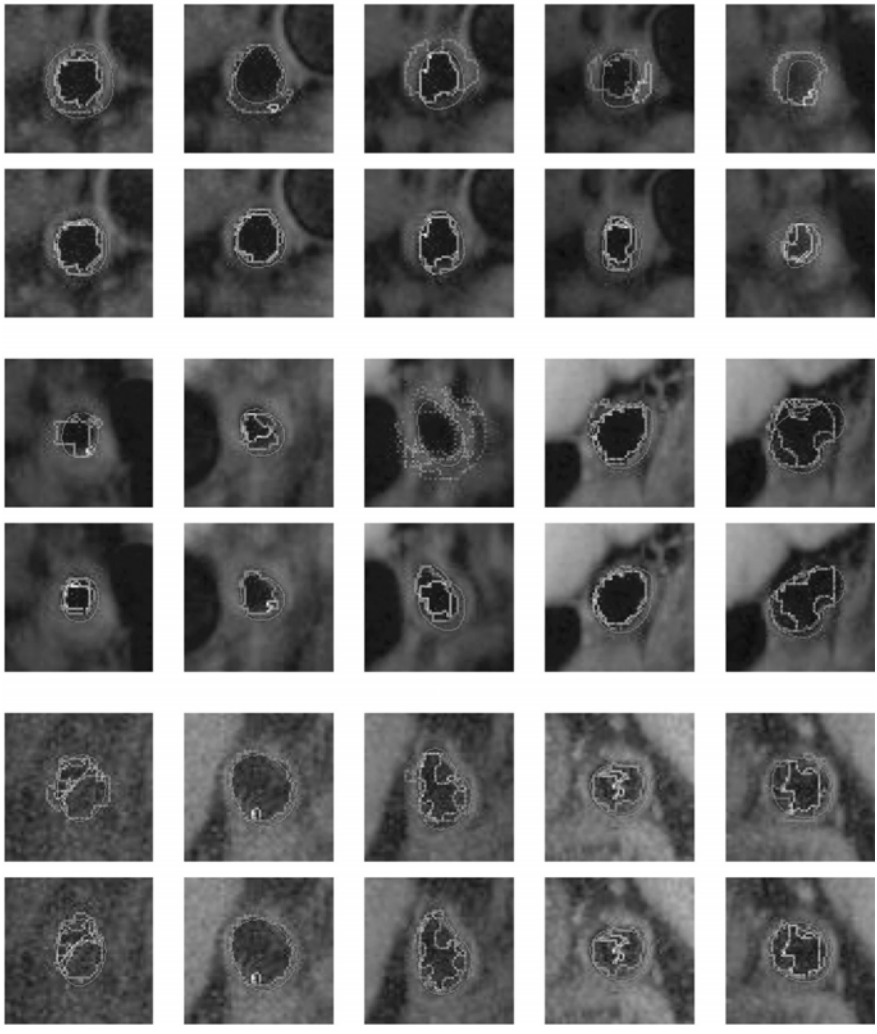


Figure 9.45: Results of estimated boundary using circular- vs. elliptical-based methods. The system used was FCM based. Top rows are circular-based ROI, while the corresponding bottom rows are elliptical-based ROIs.

9.6.4 Visualization of Circular versus Elliptical Methods

Figures 9.45–9.48 show the comparision between the outputs of the systems when the system uses circular ROI versus elliptical ROIs. The visualization results from the two systems (one with circular ROI vs. elliptical ROI) are shown in pairs: top row corresponds to circular ROI methodology, while bottom row corresponds to elliptical ROI. Note that the equation for computing the elliptical ROI is shown in Eq. (9.25).

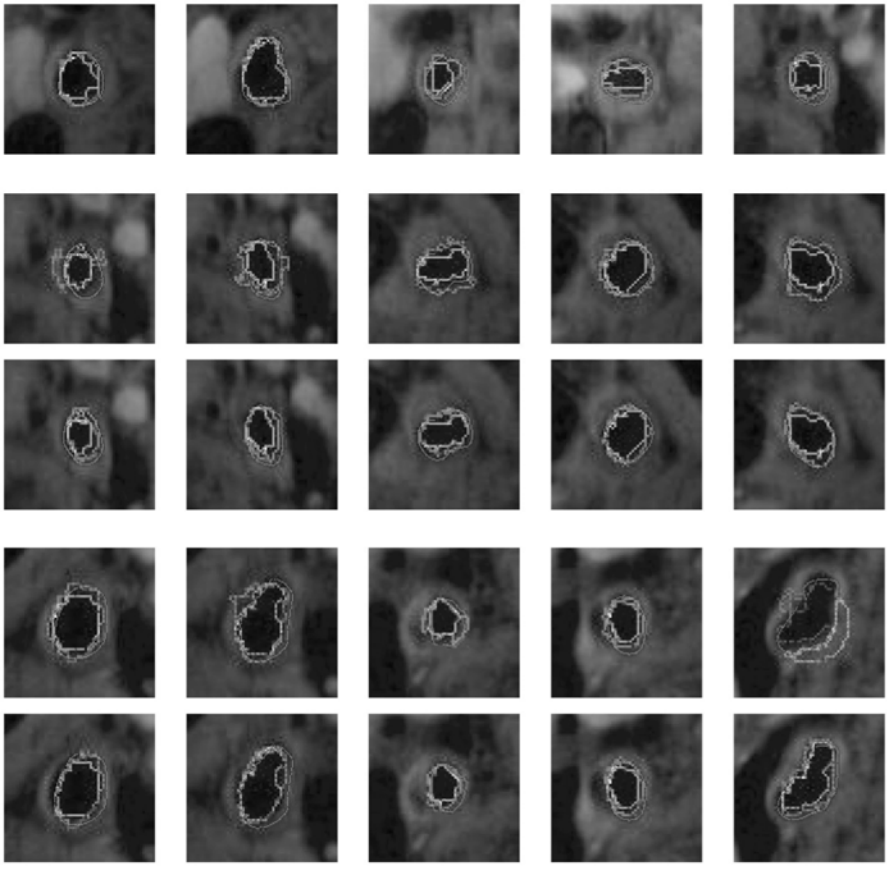


Figure 9.46: Results of estimated boundary using circular- vs. elliptical-based methods. The system used was FCM based. Top rows are circular-based ROI, while the corresponding bottom rows are elliptical-based ROIs.

9.7 Conclusions

9.7.1 System Strengths

This chapter presented the following new implementations when it comes to MR plaque imaging: (a) Application of three different sets of classifiers for lumen region classification in plaque MR protocols. These classifiers are done in multiresolution framework. Thus subregions are chosen and subclassifiers are applied to compute the accuracy of the pixel values belonging to a class. (b) Region merging for subclasses in lumen region to compute accurate lumen region and lumen boundary in cross-sectional images. (c) Rotational effect of

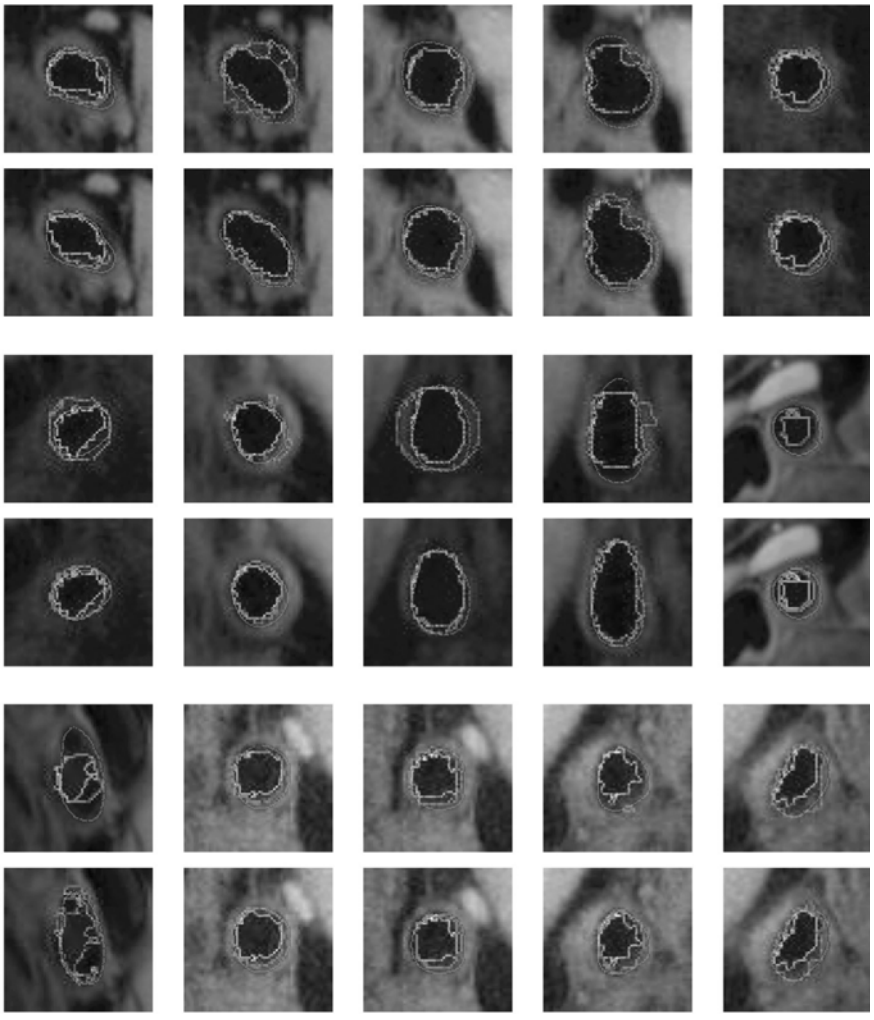


Figure 9.47: Results of estimated boundary using circular- vs. elliptical-based methods. The system used was FCM based. Top rows are circular-based ROI, while the corresponding bottom rows are elliptical-based ROIs.

region of interest in bifurcation zones for accurate lumen region identification and boundary estimation.

9.7.2 System Weakness

The ROI is determined by looking at the overlay of the ground truth contour and the grayscale image. The center is estimated by eye and the radius of the contour

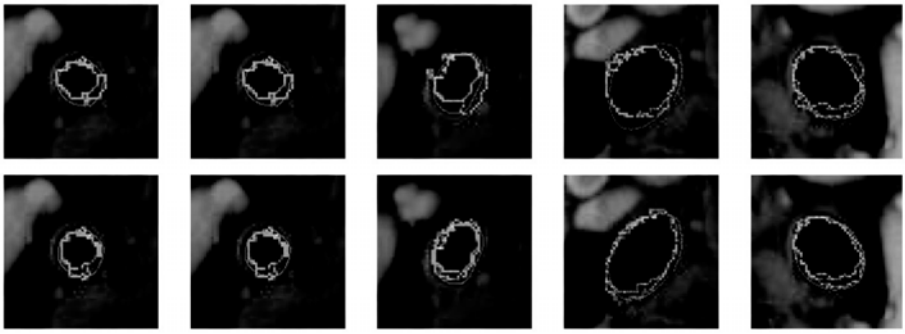


Figure 9.48: Results of estimated boundary using circular- vs. elliptical-based methods. The system used was FCM based. Top rows are circular-based ROI, while the corresponding bottom rows are elliptical-based ROIs.

in most cases is the farthest distance from the center to a point on the contour. The ROI is the circle given by this center and this radius. The center and radius are sometimes adjusted after seeing the result of the pipeline's first run.

9.8 Acknowledgments

The authors thank the Department of Radiology for the MR datasets. Thanks also to the students of Biomedical Imaging Laboratory at the Department of Biomedical Engineering, Case Western Reserve University for cooperating on sharing the calibrated machines for tracing the ground truth on plaque volumes.

Questions

1. *What is arterial remodeling? (*Lancet, *Vol. 353, pp. SII5–SII9, 1999)*
2. *What are the main challenges in lumen quantification process?*
3. *Discuss the three types of algorithms used in this chapter for lumen estimation?*
4. *What brings the low error and why?*
5. *Compare the error performance using three different systems?*

Bibliography

[1] Rogers, W. J., Prichard, J. W., Hu, Y. L., Olson, P. R., Benckart, D. H., Kramer, C. M., Vido, D. A., and Reichek, N., Characterization of signal properties in atherosclerotic plaque components by intravascular MRI, Arterioscler. Thromb. Vasc. Biol., Vol. 20, No. 7, pp. 1824–1830, 2000.

[2] Ross, R., Atherosclerosis—An inflammatory disease, N. Engl. J. Med., Vol. 340, No. 2, pp. 115–126, 1999.

[3] Reo, N. V. and Adinehzadeh, M., NMR spectroscopic analyses of liver phosphatidylcholine and phosphatidylethanolamine biosynthesis in rats exposed to peroxisome proliferators—A class of nongenotoxic hepatocarcinogens, Toxicol. Appl. Pharmacol., Vol. 164, No. 2, pp. 113–126, 2000.

[4] Pietrzyk, U., Herholz, K., and Heiss, W. D., Three-dimensional alignment of functional and morphological tomograms, J. Comput. Assist. Tomogr., Vol. 14, No. 1, pp. 51–59, 1990.

[5] Coombs, B. D., Rapp, J. H., Ursell, P. C., Reily, L. M., and Saloner, D., Structure of plaque at carotid bifurcation: High-resolution MRI with histological correlation, stroke, Vol. 32, No. 11, pp. 2516–2521, 2001.

[6] Brown, B. G., Hillger, L., Zhao, X. Q., Poulin, D., and Albers, J. J., Types of changes in coronary stenosis severity and their relative importance in overall progression and regression of coronary disease: Observations from the FATS TRial: Familial Atherosclerosis Treatement Study, Ann. N.Y. Acad. Sci., Vol. 748, pp. 407–417, 1995.

[7] Helft, G., Worthley, S. G., Fuster, V., Fayad, Z. A., Zaman, A. G., Corti, R., Fallon, J. T., and Badimon, J. J., Progression and regression of atherosclerotic lesions: Monitoring with serial noninvasive MRI, Circulation, Vol. 105, pp. 993–998, 2002.

[8] Hayes, C. E., Hattes, N., and Roemer, P. B., Volume imaging with MR phased arrays, Magn. Reson. Med., Vol. 18, No. 2, pp. 309–319, 1991.

[9] Gill, J. D., Ladak, H. M., Steinman, D. A., and Fenster, A., Segmentation of ulcerated plaque: A semi-automatic method for tracking the progression of carotid atherosclerosis, In: Proceedings of 22nd Annual EMBS International Conference, 2000, pp. 669–672.

[10] Yang, F., Holzapfel, G., Schulze-Bauer, Ch. A. J., Stollberger, R., Thedens, D., Bolinger, L., Stolpen, A., and Sonka, M., Segmentation of wall and plaque in in vitro vascular MR images, Int. J. Cardiovasc. Imaging, Vol. 19, No. 5, pp. 419–428, 2003.

[11] Kim, W. Y., Stuber, M., Boernert, P., Kissinger, K. V., Manning, W. J., and Botnar, R. M., Three-dimensional black-blood cardiac magnetic resonance coronary vessel wall imaging detects positive arterial remodeling in patients with nonsignificant coronary artery disease, Circulation, Vol. 106, No. 3, pp. 296–299, 2002.

[12] Wilhjelm, J. E., Jespersen, S. K., Hansen, J. U., Brandt, T., Gammelmark, K., and Sillesen, H., *In vitro* imaging of the carotid artery with spatial compound imaging, In: Proceedings of the 3rd Meeting of Basic Technical Research of the Japan Society of Ultrasonic in Medicine, 1999, Vol. 15, pp. 9–14.

[13] Jespersen, S. K., Gro/nholdt, M.-L. M., Wilhjelm, J. E., Wiebe, B., Hansen, L. K., and Sillesen, H., Correlation between ultrasound B-mode images of carotid plaque and histological examination, IEEE Proc. Ultrason. Symp., Vol. 2, pp. 165–168, 1996.

[14] Quick, H. H., Debatin, J. F., and Ladd, M. E., MR imaging of the vessel wall, Euro. Radiol., Vol. 12, No. 4, pp. 889–900, 2002.

[15] Corti, R., Fayad, Z. A., Fuster, V., Worthley, S. G., Helft, G., Chesebro, J., Mercuri, M., and Badimon, J. J., Effects of lipid-lowering by simvastatin on human atherosclerotic lesions: A longitudinal study by high-resolution, noninvasive magnetic resonance imaging, Circulation, Vol. 104, No. 3, pp. 249–252, 2001.

[16] Fayad, Z. A. and Fuster, V., Characterization of atherosclerotic plaques by magnetic resonance imaging, Ann. N. Y., Acad. Sci., Vol. 902, pp. 173–186, 2000.

[17] Helft, G., Worthley, S. G., Fuster, V., Zaman, A. G., Schechter, C., Osende, J. I., Rodriguez, O. J., Fayad, Z. A., Fallon, J. T., and Badimon, J. J., Atherosclerotic aortic component quantification by noninvasive magnetic resonance imaging: An in vivo study in rabbits, J. Am. Coll. Cardiol., Vol. 37, No. 4, pp. 1149, 2001.

[18] Shinnar, M., Fallon, J. T., Wehrli, S., Levin, M., Dalmacy, D., Fayad, Z. A., Badimon, J. J., Harrington, M., Harrington, E., and Fuster, V., The diagnostic accuracy of ex vivo MRI for human atherosclerotic plaque characterization, Arterioscler. Thromb., Vasc. Biol., Vol. 19, No. 11, pp. 2756–2761, 1999.

[19] Toussaint, J. F., LaMuraglia, G. M., Southern, J. F., Fuster, V., and Kantor, H. L., Magnetic resonance images lipid, fibrous, calcified, hemorrhagic, and thrombotic components of human atherosclerosis in vivo, Circulation, Vol. 94, No. 5, pp. 932–938, 1996.

[20] Worthley, S. G., Helft, G., Fuster, V., Fayad, Z. A., Rodriguez, O. J., Zaman, A. G., Fallon, J. T., and Badimon, J. J., Noninvasive in vivo magnetic resonance imaging of experimental coronary artery lesions in a porcine model, Circulation, Vol. 101, No. 25, pp. 2956–2961, 2000.

[21] Toussaint, J. F., NMR sequences for biochemical analysis and imaging of vascular diseases, Int. J. Cardiovasc. Imaging, Vol. 17, No. 3, pp. 187–194, 2001.

[22] Naghavi, M., Libby, P., Falk, E., Casscells, S. W., Litovsky, S., Rumberger, J., Badimon, J. J., Stefanadis, C., Moreno, P., Pasterkamp, G., Fayad, Z., Stone, P. H., Waxman, S., Raggi, P., Madjid, M., Zarrabi, A., Burke, A., Yuan, C., Fitzgerald, P. J., Siscovick, D. S., de Korte, C. L., Aikawa, M., Airaksinen, K. E., Assmann, G., Becker, C. R., Chesebro, J. H., Farb, A., Galis, Z. S., Jackson, C., Jang, I.-K., Koening, W., Lodder, R. A., March, K., Demirovic, J., Navab, M., Priori, S. G., Rekhter, M. D., Bahr, R., Grundy, S. M., Mehran, R., Colombo, A., Boerwinkle, E., Ballantyne, C., Insull, W., Jr., Schwartz, R. S., Vogel, R., Serruys, P. W., Hansson, G. K., Faxon, D. P., Kaul, S., Drexler, H., Greenland, P., Muller, J. E., Virmani, R., Ridker, P. M., Zipes, D. P.,

Shah, P. K., and Willerson, J. T., From vulnerable plaque to vulnerable patient: A call for new definitions and risk assessment strategies, Part I, Circulation, Vol. 108, No. 14, pp. 1662–1772, 2003.

[23] Naghavi, M., Libby, P., Falk, E., Casscells, S. W., Litovsky, S., Rumberger, J., Badimon, J. J., Stefanadis, C., Moreno, P., Pasterkamp, G., Fayad, Z., Stone, P. H., Waxman, S., Raggi, P., Madjid, M., Zarrabi, A., Burke, A., Yuan, C., Fitzgerald, P. J., Siscovick, D. S., de Korte, C. L., Aikawa, M., Airaksinen, K. E., Assmann, G., Becker, C. R., Chesebro, J. H., Farb, A., Galis, Z. S., Jackson, C., Jang, I.-K., Koenig, W., Lodder, R. A., March, K., Demirovic, J., Navab, M., Priori, S. G., Rekhter, M. D., Bahr, R., Grundy, S. M., Mehran, R., Colombo, A., Boerwinkle, E., Ballantyne, C., Insull, W., Jr., Schwartz, R. S., Vogel, R., Serruys, P. W., Hansson, G. K., Faxon, D. P., Kaul, S., Drexler, H., Greenland, P., Muller, J. E., Virmani, R., Ridker, P. M., Zipes, D. P., Shah, P. K., and Willerson, J. T., From vulnerable plaque to vulnerable patient: A call for new deinitions and risk assessment strategies, Part II, Circulation, Vol. 108, No. 15, pp. 1772–1778, 2003.

[24] Fayad, Z. A., Fuster, V., Nikolaou, K., and Becker, C., Computed tomography and magnetic resonance imaging for noninvasive coronary angiography and plaque imaging: Current and potential future concepts, Circulation, Vol. 106, No. 15, pp. 2026–2034, 2002.

[25] Fuster, V., Fayad, Z. A., and Badimon, J. J., Acute coronary syndromes: Biology, Lancet, Vol. 353, pp. SII5–SII9, 1999.

[26] Cai, J. M., Hatsukami, T. S., Ferguson, M. S., Small, R., Polissar, N. L., and Yuan, C., Classification of human carotid atherosclerotic lesions with in vivo multicontrast magnetic resonance imaging, Circulation, Vol. 106, No. 11, pp. 1368–1373 2002.

[27] Xu, D., Hwang, J.-N., and Yuan, C., Atherosclerotic blood vessel tracking and lumen segmentation in topology changes situations of MR image sequences, In: Proceedings of the International Conference on Image Processing (ICIP), 2000, Vol. 1, pp. 637–640.

[28] Xu, D., Hwang, J.-N., and Yuan, C., Atherosclerotic plaque segmentation at human carotid artery based on multiple contrast weighting

MR images, In: Proceedings of the International Conference on Image Processing (ICIP), 2001, Vol. 2, pp. 849–852.

[29] Hatsukami, T. S., Ross, R., Polissar, N. L., and Yuan, C., Visalization of fibrous cap thickness and rupture in human atherosclerotic carotid plaque in vivo with high-resolution magnetic resonance imaging, Circulation, Vol. 102, No. 9, pp. 959–964, 2000.

[30] Yuan, C., Lin, E., Millard, J., and Hwang, J. N., Closed contour edge detection of blood vessel lumen and outer wall boundaries in black-blood MR images, Magn. Reson. Imaging, Vol. 17, No. 2, pp. 257–266, 1999.

[31] Yuan, C., Kerwin, W. S., Ferguson, M. S., Polissar, N., Zhang, S., Cai, J., and Hatsukami, T. S., Contrast-enhanced high resolution MRI for atherosclerotic carotid artery tissue characterization, J. Magn. Reson. Imaging, Vol. 15, No. 1, pp. 62–67, 2002.

[32] Yuan, C., Zhang, S.-X., Polissar, N. L., Echelard, D., Ortiz, G., Davis, J. W., Ellington, E., Ferguson, M. S., and Hatsukami, T. S., Identification of fibrous cap rupture with magnetic resonance imaging is highly associated with recent transient ischemic attack or stroke, Circulation, Vol. 105, No. 2, pp. 181–185, 2002.

[33] Zhang, S., Hatsukami, T. S., Polissar, N. L., Han, C., and Yuan, C., Comparison of carotid vessel wall area measurements using three different contrast-weighted black blood MR imaging techniques, Magn. Reson. Imaging, Vol. 19, No. 6, pp. 795–802, 2001.

[34] Zhao, X. Q., Yuan, C., Hatsukami, T. S., Frechette, E. H., Kang, X. J., Maravilla, K. R., and Brown, B. G., Effects of prolonged intensive lipid-lowering therapy on the characteristics of carotid atherosclerotic plaques in vivo by MRI: A case-control study, Arterioscler. Thromb. Vasc. Biol., Vol. 21, No. 10, pp. 1623–1629, 2001.

[35] Kerwin, W. S., Han, C., Chu, B., Xu, D., Luo, Y., Hwang, J.-N., Hatsukami, T. S., and Yuan, C., A quantitative vascular analysis system for evaluation of atherosclerotic lesions by MRI, In: Proceedings of the International Conference on Medical Image Computing

and Computer-Assisted Intervention (MICCAI), 2001, Vol. 2208, pp. 786–794.

[36] Han, C., Hatsukami, T. S., and Yuan, C., A multi-scale method for automatic correction of intensity non-uniformity in MR images, J. Magn. Reson. Imaging, Vol. 13, No. 3, pp. 428–436, 2001.

[37] Zhang, Q., Wendt, M., Aschoff, A. J., Lewin, J. S., and Duerk, J. L., A multielement RF coil for MRI guidance of interventional devices, J. Magn. Reson. Imaging, Vol. 14, No. 1, pp. 56–62, 2001.

[38] Goldin, J. G., Yoon, H. C., Greaser, L. E., III, *et al.*, Spiral versus electron-beam CT for coronary artery calcium scoring, Radiology, Vol. 221, pp. 213–221, 2001.

[39] Becker, C. R., Kleffel, T., Crispin, A., *et al.*, Coronary artery calcium measurement: Agreement of multirow detector and electron beam CT, Am. J. Roentgenol., Vol. 176, pp. 1295–1298, 2001.

[40] Gaylord, G. M., Computed tomographic and magnetic resonance coronary angiography: Are you ready?, Radiol. Manag., Vol. 24, pp. 16–20, 2002.

[41] Haberl, R., Becker, A., Leber, A., *et al.*, Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: Results of 1,764 patients, J. Am. Coll. Cardio., Vol. 37, pp. 451–457, 2001.

[42] Leber, A. W., Knez, A., Mukherjee, R., *et al.*, Usefulness of calcium scoring using electron beam computed tomography and noninvasive coronary angiography in patients with suspected coronary artery disease, Am. J. Cardiol., Vol. 88, pp. 219–223, 2001.

[43] McConnell, M. V., Imaging techniques to predict cardiovascular risk, Curr. Cardiol. Rep., Vol. 2, pp. 300–307, 2000.

[44] Ohnesorge, B., Flohr, T., Fischbach, R., *et al.*, Reproducibility of coronary calcium quantification in repeat examinations with retrospectively ECG-gated multisection spiral CT, Euro. Radiol., Vol. 12, pp. 1532–1540, 2002.

[45] Sevrukov, A., Jelnin, V., and Kondos, G. T., Electron beam CT of the coronary arteries: Cross-sectional anatomy for calcium scoring, Am. J. Roentgenol., Vol. 177, pp. 1437–1445, 2001.

[46] Bond, J. H., Colorectal cancer screening: The potential role of virtual colonoscopy, J. Gastroenterol., Vol. 37, No. 13, pp. 92–96, 2002.

[47] Chaoui, A. S., Blake, M. A., Barish, M. A., *et al.*, Virtual colonoscopy and colorectal cancer screening, Abdom. Imaging, Vol. 25, pp. 361–367, 2000.

[48] Dobos, N. and Rubesin, S. E., Radiologic imaging modalities in the diagnosis and management of colorectal cancer, Hematol. Oncol. Clin. N. Am., Vol. 16, No. X, pp. 875–895, 2002.

[49] Fenlon, H. M., Nunes, D. P., Clarke, P. D., *et al.*, Colorectal neoplasm detection using virtual colonoscopy: A feasibility study, Gut, Vol. 43, pp. 806–811, 1998.

[50] Fenlon, H. M., Nunes, D. P., Schroy, P. C., III, *et al.*, A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps, N. Engl. J. Med., Vol. 341, pp. 1496–1503, 1999.

[51] Ferrucci, J. T., Colon cancer screening with virtual colonoscopy: Promise, polyps, politics, Am. J. Roentgenol., Vol. 177, pp. 975–988, 2001.

[52] Harvey, C. J., Renfrew, I., Taylor, S., *et al.*, Spiral CT pneumocolon: Applications, status and limitations, Euro. Radiol., Vol. 11, pp. 1612–1625, 2001.

[53] Mendelson, R. M., Foster, N. M., Edwards, J. T., *et al.*, Virtual colonoscopy compared with conventional colonoscopy: A developing technology, Med. J. Aust., Vol. 173, pp. 472–475, 2000.

[54] Rex, D. K., Considering virtual colonoscopy, Rev. Gastroenterol. Disord. Vol. 2, pp. 97–105, 2002.

[55] Xynopoulos, D., Stasinopoulou, M., Dimitroulopoulos, D., *et al.*, Colorectal polyp detection with virtual colonoscopy (computed tomographic colonography): The reliability of the method, Hepatogastroenterology, Vol. 49, No. 43, pp. 124–127, 2002.

[56] Budoff, M. J., Oudiz, R. J., Zalace, C. P., *et al.*, Intravenous three dimensional coronary angiography using contrast enhanced electron beam computed tomography, Am. J. Cardiol., Vol. 83, pp. 840–845, 1999.

[57] Lu, B., Budoff, M. J., Zhuang, N., *et al.*, Causes of interscan variability of coronary artery calcium measurements at electron-beam CT, Acad. Radiol., Vol. 9, pp. 654–661, 2002.

[58] Mao, S., Budoff, M. J., Bakhsheshi, H., Liu, S. C., Improved reproducibility of coronary artery calcium scoring by electron beam tomography with a new electrocardiographic trigger method, Invest. Radiol., Vol. 36, pp. 363–367, 2001.

[59] Bielak, L. F., Rumberger, J. A., Sheedy, P. F., II, *et al.*, Probabilistic model for prediction of angiographically defined obstructive coronary artery disease using electron beam computed tomography calcium score strata, Circulation, Vol. 102, No. 4, pp. 380–385, 2000.

[60] Takahashi, N. and Bae, K. T., Quantification of coronary artery calcium with multi-detector row CT: Assessing interscan variability with different tube currents-pilot study, Radiology, Vol. 228, No. 1, pp. 101–106, 2003.

[61] Pannu, H. K., Flohr, T. G., Corl, F. M., and Fishman, E. K., Current concepts in multi-detector row CT evaluation of the coronary arteries: Principles, Techniques, and Anatomy, Radiographics, Vol. 23, No. 90001, pp. S111–S125, 2003.

[62] Suri, J. S. and Liu, K., A review on MR vascular image processing algorithms: Acquisition and pre-filtering, Part-I, IEEE Trans. Inf. Tech. Biomed., Vol. 6, 2001.

[63] Suri, J. S., A review on MR vascular image processing algorithms: Skeleton vs. Non-skeleton Approaches, Part-II, IEEE Trans. Inf. Technol. Biomed., Vol. 6, 2002.

[64] Suri, J. S., An algorithm for time-of-flight black blood vessel detection. In: Proceedings of the 4th IASTED International Conference in Signal Processing, 2002, pp. 560–564.

[65] Suri, J. S., Artery–Vein detection in very noisy TOF angiographic volume using dynamic feedback scale-space ellipsoidal filtering, In: Proceedings of the 4th IASTED International Conference in Signal Processing, 2002, pp. 565–571.

[66] Suri, J. S. and Laxminarayan, S. N., Angiography and Plaque Imaging: Advances Segmentation Techniques, CRC Press, Boca Raton, FL, 2003.

[67] Suri, J. S., Wilson, D. L., and Laxminarayan, S. N., Handbook of Medical Image Analysis: Segmentation and Registration Models, Marcel Dekker, New York, 2004.

[68] Suri, J. S., Two automatic training-based forced calibration algorithms for left ventricle boundary estimation in cardiac images, Int. Conf. IEEE Eng. Med. Biol., Vol. 2, pp. 538–541, 1997.

[69] Suri, J. S., Computer vision, pattern recognition and image processing in left ventricle segmentation: The last 50 years, Int. J. Patt. Anal. Appl., Vol. 3, No. 3, pp. 209–242, 2000.

[70] Suri, J. S., Kamaledin, S., and Singh, S., Advanced Algorithmic Approaches to Medical Image Segmentation: State-of-the-Art Applications in Cardiology, Neurology, Mammography and Pathology, 2001.

[71] Suri, J. S. and Laxminarayan, S. N., PDE and Level Sets: Algorithmic Approaches to Static and Motion Imagery, Kluwer Academic/Plenum Publishers, 2002.

[72] Salvado, O., Hillenbrand, C., Zhang, S., Suri, J. S., and Wilson, D., MR signal inhomogeneity correction for visual and computerized atherosclerosis lesion assessment, In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Arlington, VA, 2004.

[73] Yabushita, H., Bouma, B. E., Houser, S. L., Aretz, H. T., Jang, I.-K., Schlendorf, K. H., Kauffman, C. R., Shishkov, M., Kang, D.-H., Halpern, E. F., and Tearney, G. J., Characterization of human atherosclerosis by optical coherence tomography, Circulation, Vol. 106, No. 13, pp. 1640–1645, 2002.

[74] Nair, A., Kuban, B. D., Obuchowski, N., and Vince, D. G., Assessing spectral algorithms to predict atherosclerotic plaque composition with normalized and raw intravascular ultrasound data, Ultrasound Med. Biol., Vol. 27, No. 10, pp. 1319–1331, 2001.

[75] Wink, O., Niessen, W. J., and Viergever, M. A., Fast delineation and visualization of vessels in 3-D angiographic images, IEEE Trans. Med. Imaging, Vol. 19, No. 4, pp. 337–346, 2000.

[76] Wink, O., Niessen, W. J., and Viergever, M. A., Fast quantification of abdominal aortic aneurysms from CTA volumes, In: Proceedings of Medical Image Computing and Computer Assisted Intervention, 1998, pp. 138–145.

[77] Niessen, W. J., Montauban van Swijndregt, A. D., Elsman, B. H. P., Wink, O., Mali, W. P. Th. M., and Viergever, M. A., Improved artery visualization in blood pool MRA of the peripheral vasculature, In: Proceedings on Computer Assisted Radiology and Surgery (CARS), 1999, pp. 119–123.

[78] Udupa, J. K., Odhner, D., Tian, J., Holland, G., and Axel, L., Automatic clutter free volume rendering for MR angiography using fuzzy connectedness, SPIE Proc., Vol. 3034, pp. 114–119, 1997.

[79] Saha, P. K. and Udupa, J. K., Scale-based fuzzy connectivity: A novel image segmentation methodology and its validation, Proc. SPIE Conf. Med. Imaging, Vol. 3661, pp. 246–257, 1999.

[80] Udupa, J. K. and Samarasekera, S., Fuzzy connectedness and object delineation: Theory, algorithm, and applications in image segmentation, Graph. Models Image Process., Vol. 58, No. 3, pp. 246–261, 1996.

[81] Saha, P. K., Udupa, J. K., and Odhner, D., Scale-based fuzzy connected image segmentation: theory, algorithm, and validation, Comput Vis. Image Understanding, Vol. 77, No. 2, pp. 145–174, 2000.

[82] Udupa, J. K. and Odhner, D., Shell rendering, IEEE Comput Graph. Appl., Vol. 13, No. 6, pp. 58–67, 1993.

[83] Lei, T., Udupa, J. K., Saha, P. K., and Odhner, D., MR angiographic visualization and artery-vein separation, Proc. of SPIE, Int. Soc. Opt. Eng., Vol. 3658, pp. 58–66, 1999.

[84] Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., and Kikinis, R., Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images, Med. Image Anal., Vol. 2, No. 2, pp. 143–168, 1998.

[85] Sato, Y., Chen, J., Harada, N., Tamura, S., and Shiga, T., Automatic extraction and measurements of leukocyte motion in micro vessels using spatiotemporal image analysis, IEEE Trans. Biomed. Eng., Vol. 44, No. 4, pp. 225–236, 1997.

[86] Sato, Y., Nakajima, S., Atsumi, H., Koller, T., Gerig, G, Yoshida, S., and Kikinis, R., 3-D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images, In: Proceedings on CVRMed and MRCAS (CVRMed/MRCAS), 1997, pp. 213–222.

[87] Sato, Y., Araki, T., Hanayama, M., Naito, H., and Tamura, S., A viewpoint determination system for stenosis diagnosis and quantification in coronary angiographic image acquisition, IEEE Trans. Med. Imaging, Vol. 17, No. 1, pp. 121–137, 1998.

[88] Frangi, A. F., Niessen, W. J., Hoogeveen, R. M., van Walsum, Th., and Viergever, M. A., Model-based quantification of 3-D magnetic resonance angiographic images, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 946–956, 1999.

[89] Berliner, J. A., Navab, M., Fogelman, A. M., Frank, J. S., Demer, L. L., Edwards, P. A., Watson, A. D., and Lusis, A. J., Atherosclerosis: Basic mechanismsm Oxidation, inflammation, and genetics, Circulation, Vol. 91, No. 9, pp. 2488–2496, 1995.

[90] Botnar, R. M., Stuber, M., Kissinger, K. V., Kim, W. Y., Spuentrup, E., and Manning, W. J., Noninvasive coronary vessel wall and plaque imaging with magnetic resonance imaging, Circulation, Vol. 102, No. 21, pp. 2582–2587, 2000.

[91] Breen, M. S., Lancaster, T. L., Lazebnik, R., Aschoff, A. J., Nour S. G., Lewin J. S., and Wilson, D. L., Three dimensional correlation of MR images to muscle tissue response for interventional MRI thermal ablation, Proc. SPIE Med. Imaging, Vol. 5029, pp. 202–209, 2001.

[92] Carrillo, A., Wilson, D. L., Duerk, J. L., and Lewin, J. S., Semi-automatic 3D image registration and applied to interventional MRI liver cancer treatment, IEEE Trans. Med. Imaging, Vol. 19, No. 3, pp. 175–185, 2003.

[93] Chalan, V. and Kim, Y., A methodology for evaluation of boundary detection algorithms on medical images, IEEE Trans. Med. Imaging, Vol. 16, No. 5, pp. 642–652, 1997.

[94] Clarke, L. P., Velthuizen, R. P., Camacho, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., and Silbiger, M. L., MRI segmentation: Methods and applications, Magn. Reson. Imaging, Vol. 13, No. 3, pp. 343–368, 1995.

[95] Correia, L. C. L., Atalar, E., Kelemen, M. D., Ocali, O., Hutchins, G. M., Fleg, J. L., Gerstenblith, G., Zerhouni, E. A., and Lima, J. A. C., Intravascular magnetic resonance imaging of aortic atherosclerotic plaque composition, Arterioscler. Thromb. Vasc. Biol., Vol. 17, No. 12, pp. 3626–3632, 1997.

[96] Hagberg, G., From magnetic resonance spectroscopy to classification of tumors: A review of pattern recognition methods., NMR Biomed., Vol. 11, No. 4/5, p. 148, 1998.

[97] Hajnal, J. V., Saeed, N., Soar, E. J., Oatridge, A., Young, I. R., and Bydder, G., A registration and interpolation procedure for subvoxel matching of serially acquired MR images, J. Comput. Assist. Tomography, Vol. 19, No. 2, pp. 289–296, 1995.

[98] Hurst, G. C., Hua, J., Duerk, J. L., and Cohen, A. M., Intravascular (catheter) NMR receiver probe: Preliminary design analysis and application to canine iliofemoral imaging, Magn. Reson. Med., Vol. 24, No. 2, p. 343, 1992.

[99] Klingensmith, J. D., Shekhar, R., and Vince, D. G., Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images, IEEE Trans. Med. Imaging, Vol. 19, No. 10, pp. 996–1011, 2000.

[100] Ladak, H. M., Thomas, J. B., Mitchell, J. R., Rutt, B. K., and Steinman, D. A., A semi-automatic technique for measurement of arterial wall from black blood MRI, Med. Phy., Vol. 28, No. 6, p. 1098, 2001.

[101] Lancaster, T. L. and Wilson, D. L., Correcting spatial distortion of histological images, Ann. Biomed. Eng., Vol. XX, No. X, pp. XXX-XXX, 2003.

[102] Lazebnik, R., Lancaster, T. L., Breen, M. S., Lewin J. S., and Wilson, D. L., Volume registration using needle paths and point landmarks for evaluation of interventional MRI treatments, IEEE Trans. Med. Imaging, Vol. 22, No. 5, pp. 659–660, 2003.

[103] Lorigo, L. M., Faugeras, O., Grimson, W. E. L., Keriven, R., Kikinis, R., Nabavi, A., and Westin, C. F., Codimension-two geodesic active contours for the segmentation of tubular structures, Vol. 1, No. 13–15, pp. 444–451, 2000.

[104] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximization of mutual information, IEEE Trans. Med. Imaging, Vol. 16, No. 2, pp. 187–198, 1997.

[105] Merickel, M. B., Carman, C. S., Watterson, W. K., Brookeman, J. R., and Ayers, C. R., Multispectral pattern recognition of MR imagery for the noninvasive analysis of atherosclerosis, In: 9th International Conference on Pattern Recognition, 1988, pp. 1192–1197.

[106] Pallotta, S., Gilardi, M. C., Bettinardi, B., Rizzo, G., Landoni, C., Striano, G., Masi, R., and Fazio, F., Application of a surface matching image registration technique to the correlation of cardiac studies in position emission tomography (PET) by transmission images, Phy. Med. Biol., Vol. 40, No. 10, pp. 1695–1708, 1995.

[107] Pelizzari, C. A., Chen, G. T. Y., Spelbring, D. R., Weichselbaum, R. R., and Chen, C. T., Accurate three-dimensional registration of CT, PET

and/or MR images of the brain, J. Comput. Assist. Tomography, Vol. 13, No. 1, pp. 20–26, 1989.

[108] Pietrzyk, U., Herholz, K., Fink, G., Jacobs, A., Mielke, R., Slansky, I., Michael, W., and Heiss, W. D., An interactive tecnique for three-dimensional image registration: Validation for PET, SPECT, MRI and CT brain studies, J. Nuclear Med., Vol. 35, No. 12, pp. 2011–2018, 1994.

[109] Rioufol, G., Finet, G., Ginon, I., Andre, F., XRossi, R., Vialle, E., Desjoyaux, E., Convert, G., Huret, J. F., and Tabib, A., Multiple atherosclerotic plaque rupture in acute coronary syndrome: A three-vessel intravascular ultrasound study, Circulation, Vol. 106, No. 7, p. 804, 2002.

[110] Saeed, N., Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation, NMR Biomed., Vol. 11, No. 4/5, pp. 157, 1998.

[111] Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A., and Leahy, R. M., Magnetic resonance image tissue classification using a partial volume model, Neuroimage, Vol. 13, No. 5, p. 856, 2001.

[112] Thieme, T., Wernecke, K. D., Meyer, R., Brandenstein, E., Habedank, D., Hinz, A., Felix, S. B., Baumann, G., and Kleber, F. X., Angioscopic evaluation of atherosclerotic plaques: Validation by histomorphologic analysis and association with stable and unstable coronary syndromes, J. Am. Coll. Cardiol., Vol. 28, No. 1, pp. 1–6, 1996.

[113] Trouard, T. P., Altbach, M. I., Hunter, G. C., Eskelson, C. D., and Gmitro, A. F., MRI and NMR spectroscopy of the lipids of atherosclerotic plaque in rabbits and humans, Magn. Reson. Med., Vol. 38, No. 1, pp. 19–26, 1997.

[114] van den Elsen, P. A., Pol, E. J. D., and Viergever, M. A., Medical image matching—A review with classification, IEEE Eng. Med. Biol., Vol. 12, No. 1, pp. 26–39, 1993.

[115] Viola, P. A. and Wells, W. M., III, Alignment by maximization of mutual information, In: IEEE Proceedings of the 5th International Conference on Computer Vision, 1995, pp. 16–23.

[116] Weber, D. A. and Ivanovic, M., Correlative image registration, Semin. Nuclear Med., Vol. 24, No. 4, pp. 311–323, 1994.

[117] West, J., Fitzpatrick, M., Wang, M. Y., Dawant, B. M., Maurer, C. R., Kessler, M. L., Maciunas, R. J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P. A., Napel, S., Sumanaweera, T. S., Harkness, B. A., Hemler, P. F., Hill, D. L. G., Hawkes, D. J., Studholme, C., Maintz, J. B., Viergever, M. A., Malandain, G., Pennec, X., Noz, M. E., Maguire, G. Q., Pollack, M., Pelizzari, C. A., Robb, R. A., Hanson, D., and Woods, R. P., Comparison and evaluation of retrospective intermodality brain image registration techniques, J. Comput. Assist. Tomography, Vol. 21, No. 4, pp. 554–566, 1997.

[118] Breen, M. S., Lancaster T. L., Lazebnik, R., Nour S. G., Lewin J. S., and Wilson, D. L., Three dimensional method for comparing in vivo interventional MR images of thermally ablated tissue with tissue response, J. Magn. Reson. Imaging, Vol. 18, No. 1, pp. 90–102, 2003.

[119] Wilson, D. L., Carrillo, A., Zheng, L., Genc, A., Duerk, J. L., and Lewin, J. S., Evaluation of 3D image registration as applied to MR-guided thermal treatment of liver cancer, J. Magn. Reson. Imaging, Vol. 8, No. 1, pp. 77–84, 1998.

[120] Wink, O., Fast delineation and visualization of vessels in 3-D angiographic images, IEEE Trans. Med. Imaging, Vol. 19, No. 4, pp. 337–346, 2000.

[121] Yu, J. N., Fahey, F. H., Gage, H. D., Eades, C. G., Harkness, B. A., Pelizzari, C. A., and Keyes, J. W., Intermodality, retrospective image registration in the thorax, J. Nuclear Med., Vol. 36, No. 12, pp. 2333–2338, 1995.

[122] Draney, M. T., Herfkens, R. J., Hughes, T. J. R., Plec, N. J., Wedding, K. L., Zarins, C. K., and Taylor, C. A., Quantification of vessel wall cyclic strain using cine phase contrast magnetic resonance imaging, Ann. Biomed. Eng., Vol. 30, No. 8, pp. 1033–1045, 2002.

[123] MacNeill, B. D., Lowe, H. C., Takano, M., Fuster, V., and Jang, I.-K., Intravascular modalities for detection of vulnerable plaque: current

status, Arterioscler. Thromb. Vasc. Biol., Vol. 23, No. 8, pp. 1333–1342, 2003.

[124] Ziada, K., Tuzcu, E. M., Nissen, S. E., Ellis, S., Whitlow, P. L., and Franco, I., Prognostic importance of various intravascular ultrasound measurements of lumen size following coronary stenting (submitted).

[125] Ziada, K., Kapadia, S., Tuzcu, E. M., and Nissen, S. E., The current status of intravascular ultrasound imaging, Curr. Prob. Cardiol., Vol. 24, No. 9, pp. 541–616, 1999.

[126] Nair, A., Kuban, B. D., Tuzcu, E. M., Schoenhagen, P., Nissen, S. E., and Vince, D. G., Coronary plaque classification with intravascular ultrasound radiofrequency data analysis, Circulation, Vol. 106, No. 17, pp. 2200–2206, 2002.

[127] Nissen, S. E. and Yock, P., Intravascular ultrasound: Novel pathophysiological insights and current clinical applications, Circulation, Vol. 103, No. 4, pp. 604–616, 2001.

[128] Woods, R. P., Cherry, S. R., and Mazziotta, J. C., Rapid automated algorithm for aligning and reslicing PET images, J. Comput. Assist. Tomography, Vol. 16, No. 4, pp. 620–633, 1992.

[129] Woods, R. P., Mazziotta, J. C., and Cherry, S. R., MRI-PET registration with automated algorithm, J. Comput. Assist. Tomography, Vol. 17, No. 4, pp. 536–546, 1993.

[130] Fei, B. W., Boll, D. T., Duerk, J. L., and Wilson, D. L., Image registration for interventional MRI-guided minimally invasive treatment of prostate cancer, In: The 2nd Joint Meeting of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, 2002, Vol. 2, p. 1185.

[131] Fei, B. W., Duerk, J. L., Boll, D. T., Lewin, J. S., and Wilson, D. L., Slice to volume registration and its potential application to interventional MRI guided radiofrequency thermal ablation of prostate cancer, IEEE Trans. Med. Imaging, Vol. 22, No. 4, pp. 515–525, 2003.

[132] Fei, B. W., Duerk, J. L., and Wilson, D. L., Automatic 3D registration for interventional MRI-guided treatment of prostate cancer, Comput. Aided Surg., Vol. 7, No. 5, pp. 257–267, 2002.

[133] Fei, B. W., Frinkley K., and Wilson, D. L., Registration algorithms for interventional MRI-guided treatment of the prostate cancer, Proc. SPIE Med. Imaging, Vol. 5029, pp. 192–201, 2003.

[134] Fei, B. W., Kemper, C., and Wilson, D. L., A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes, Comput. Med. Imaging Graph., Vol. 27, No. 4, pp. 267–281, 2003.

[135] Fei, B. W., Kemper, C., and Wilson, D. L., Three-dimensional warping registration of the pelvis and prostate, In: Proceedings of SPIE Medical Imaging on Image Processing, Sonka, M. and Fitzpatrick, J. M., eds., Vol. 4684, pp. 528–537, 2002.

[136] Fei, B. W., Wheaton, A., Lee, Z., Duerk, J. L., and Wilson, D. L., Automatic MR volume registration and its evaluation for the pelvis and prostate, Phy. Med. Biol., Vol. 47, No. 5, pp. 823–838, 2002.

[137] Fei, B. W., Wheaton, A., Lee, Z., Nagano, K., Duerk, J. L., and Wilson, D. L., Robust registration algorithm for interventional MRI guidance for thermal ablation of prostate cancer, In: Proceedings of SPIE Medical Imaging on Visualization, Display, and Image-Guided Procedures, Ki Mun, S., ed., Vol. 4319, pp. 53–60, 2001.

[138] Wilson, D. L. and Fei, B. W., Three-dimensional semiautomatic warping registration of the prostate and pelvis, Med. Phy. (submitted).

[139] Studholme, C., Hill, D. L. G., and Hawkes, D. J., Automated 3D registration of MR and CT images of the head, Med. Image Anal., Vol. 1, No. 2, pp. 163–175, 1996.

[140] Studholme, C., Hill, D. L. G., and Hawkes, D. J., Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Med. Phy., Vol. 24, No. 1, pp. 25–35, 1997.

[141] Song, C. Z. and Yuille, A., Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, IEEE Trans. Patt. Anal. Machine Intell., Vol. 18, No. 9, pp. 884–900, 1996.

[142] Stary, H. C., Chandler, A. B., Glagov, S., Guyton, J. R., Insull, W. J., Rosenfeld, M. E., Schaffer, S. A., Schwartz, C. J., Wagner, W. D., and Wissler, R. W., A definition of initial, fatty streak, and intermediate lesions of atherosclerosis: A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association, Arterioscler. Thromb., Vol. 14, No. 5, pp. 840–856, 1994.

[143] Hemler, P. F., Napel, S., Sumanaweera, T. S., Pichumani, R., van den Elsen, P. A., Martin, D., Drace, J., Adler, J. R., and Perkash, I., Registration error quantification of a surface-based multimodality image fusion system, Med. Phy., Vol. 22, No. 7, pp. 1049–1056, 1995.

[144] Suri, J. S., White matter/Gray matter boundary segmentation using geometric snakes: A fuzzy deformable model, In: International Conference in Application in Pattern Recognition (ICAPR), Rio de Janeiro, Brazil, March 11–14, 2001.

[145] Zhang, J., The mean field theory in EM procedures for Markov random fields, IEEE Trans. Signal Process., Vol. 40, No. 10, 1992.

[146] Kapur, T., Model Based Three Dimensional Medical Image Segmentation, Ph.D. Thesis, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, May 1999.

[147] Li, S., Markov Random Field Modeling in Computer Vision, Springer Verlag, Berlin, 1995. ISBN 0-387-701-451.

[148] Held, K., Rota Kopps, E., Krause, B., Wells, W., Kikinis, R., and Muller-Gartner, H., Markov random field segmentation of brain MR images, IEEE Trans. Med. Imaging, Vol. 16, No. 6, pp. 878–887, 1998.

[149] Witkin, A. P., Scale-space filtering, In: Proceedings of 8th International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany, 1983, Vol. 2, pp. 1019–1023.

[150] Koenderink, J. J., The structure of images, Biol. Cyb., Vol. 50, pp. 363–370, 1984.

[151] Koller, T. M., Gerig, G., Székely, G., and Dettwiler, D., Multiscale detection of curvilinear structures in 2-D and 3-D image data, In: IEEE International Conference on Computer Vision (ICCV), 1995, pp. 864–869.

[152] Koller, T. M., From Data to Information: Segmentation, Description and Analysis of the Cerebral Vascularity, Ph.D. Thesis, Swiss Federal Institute of Technology, Zürich, 1995.

[153] Gerig, G., Koller, M. Th., Székely, Brechbuhler, C., and Kubler, O., Symbolic description of 3-D structures applied to cerebral vessel tree obtained from MR angiography volume data, In: Proceedings of IPMI, Series Lecture Notes in Computer Science, Vol. 687, Barett, H. H. and Gmitro, A. F., eds., Springer-Verlag, Berlin, pp. 94–111, 1993.

[154] Thirion, J. P. and Gourdon, A., The 3-D marching lines algorithm, Graph. Models Image Process., Vol. 58, No. 6, pp. 503–509, 1996.

[155] Lindeberg, T., Scale-space for discrete signals, IEEE Patt. Anal. Machine Intell., Vol. 12, No. 3, pp. 234–254, 1990.

[156] Lindeberg, T., On scale selection for differential operators, In: Proceedings of the 8th Scandinavian Conference on Image Analysis (SCIA), 1993, pp. 857–866.

[157] Lindeberg, T., Detecting salient blob-like image structures and their scales with a scalespace primal sketch: A method for focus of attention, Int. J. Comput. Vision, Vol. 11, No. 3, pp. 283–318, 1993.

[158] Lindeberg, T., Edge detection and ridge detection with automatic scale selection, In: Proceedings of Computer Vision and Pattern Recognition, 1996, pp. 465–470.

[159] Alyward, S., Bullitte, E., Pizer, S., and Eberly, D., Intensity ridge and widths for tubular object segmentation and description, In: Proceedings of Workshop Mathematical Methods Biomedical Image Analysis (WMMBIA), Amini, A. A. and Bookstein, F. L., eds., pp. 131–138, 1996.

[160] Lorenz, C., Carlsen, I.-C., Buzug, T. M., Fassnacht, C., and Wesse, J., Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2-D and 3-D medical images, In: Proceedings of Joint Conference on CVRMed and MRCAS, 1997, pp. 233–242.

[161] Fidrich, M., Following features lines across scale, In: Proceedings of Scale-Space Theory in Computer Vision, Series Lectures Notes in Computer Science, Vol. 1252, ter Haar Romeny, B., Florack, L., Loeenderink, J., and Viergever, M., eds., Springer-Verlag, Berlin, pp. 140–151, 1997.

[162] Lindeberg, T., Feature detection with automatic scale-space selection, Int. J. Comput. Vision, Vol. 30, No. 2, pp. 79–116, 1998.

[163] Prinet, V., Monga, O., and Rocchisani, J. M., Vessels Representation in 2D and 3D Angiograms, International Congress Series (ICS), Vol. 1134, pp. 240–245, 1998. ISSN 0531-5131.

[164] Prinet, V., Monga, O., Ge, C., Loa, X. S., and Ma, S., Thin network extraction in 3-D images: Application of medial angiograms, In: International Conference on Pattern Recognition, Aug. 1996.

[165] Griffin, L., Colchester, A., and Robinson, G., Scale and segmentation of images using maximum gradient paths, Image Vision Comput., Vol. 10, No. 6, pp. 389–402, 1992.

[166] Koenderink, J. and van Doorn, A., Local features of smooth shapes: Ridges and course, In: SPIE Proceedings on Geometric Methods in Computer Vision-II, 1993, Vol. 2031, pp. 2–13.

[167] Koenderink, J. and van Doorn, A., Two-plus-one-dimensional differential geometry, Patt. Recogn. Lett., Vol. 15, No. 5, pp. 439–444, 1994.

[168] Majer, P., A statistical approach to feature detection and scale selection in images, Ph.D. Thesis, University of Göttingen, Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen, Germany, July 2000.

[169] Wells, W. M., III, Grimson, W. E. L., Kikinis, R., and Jolesz, F. A., Adaptive segmentation of MRI data, IEEE Trans. Med. Imaging, Vol. 15, No. 4, pp. 429–442, 1992.

[170] Gerig, G., Kubler, O., and Jolesz, F. A., Nonlinear anisotropic filtering of MRI data, IEEE Trans. Med. Imaging, Vol. 11, No. 2, pp. 221–232, 1992.

[171] Joshi, M., Cui, J., Doolittle, K., Joshi, S., Van Essen, D., Wang, L., and Miller, M. I., Brain segmentation and the generation of cortical surfaces, Neuroimage, Vol. 9, No. 5, pp. 461–476, 1999.

[172] Dempster, A. D., Laird, N. M., and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc., Vol. 39, pp. 1–37, 1977.

[173] Kao, Y.-H., Sorenson, J. A., Bahn, M. M., and Winkler, S. S., Dual-Echo MRI segmentation using vector decomposition and probability technique: A two tissue model, Magn. Reson. Med., Vol. 32, No. 3, pp. 342–357, 1994.

[174] Geman, S. and Geman, D., Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, IEEE Trans. Patt. Anal. Machine Intell., Vol. 6, pp. 721–741, 1984.

[175] Bezdek, J. C. and Hall, L. O., Review of MR image segmentation techniques using pattern recognition, Med. Phy., Vol. 20, No. 4, pp. 1033–1048, 1993.

[176] Hall, L. O. and Bensaid, A. M., A comparison of neural networks and fuzzy clustering techniques in segmenting MRI of the brain, IEEE Trans. Neural Networks, Vol. 3, No. 5, pp. 672–682, 1992.

Chapter 10

Hessian-Based Multiscale Enhancement, Description, and Quantification of Second-Order 3-D Local Structures from Medical Volume Data

Yoshinobu Sato[1]

10.1 Introduction

With high-resolution three-dimensional (3-D) imaging modalities becoming commonly available in medical imaging, a strong need has arisen for a means of accurate extraction and 3D quantification of the anatomical structures of interest from acquired volume data. Three-dimensional local structures have been shown to be useful for 3-D modeling of anatomical structures to improve their extraction and quantification [1–16]. In this chapter, we describe an approach to enhancement, description, and quantification of the anatomical structures characterized by second-order 3D local structures, that is, line, sheet, and blob structures.

The human body contains various types of line, sheet, and blob structures. For example, blood vessels, bone cortices, and nodules are characterized by line, sheet, and blob structures, respectively. We present a theoretical framework for systematic analysis of second-order local structures in volume data. A set of volume data is typically represented as a discrete set of samples on a regular grid. The basic approach is to analyze the continuous volume intensity function

[1]Division of Interdisciplinary Image Analysis, Osaka University Graduate School of Medicine, 2-2-D11 Yamada-oka, Suita, Osaka 565-0871, Japan

that underlies the discrete sample data. Second-order local structures around a point of interest in the underlying continuous function can be fully represented using up to second derivatives at the point, that is, the gradient vector and Hessian matrix. In order to reduce noise as well as deal with second-order local structures of "various sizes," isotropic Gaussian smoothing with different standard deviation (SD) values is combined with derivative computation. Combining Gaussian smoothing has another effect that accurate derivative computation of the Gaussian smoothed version of the underlying "continuous" function is possible by convolution operations within a size-limited local window.

In this chapter, the following topics are discussed:

- Multiscale enhancement filtering of second-order local structures, that is, line, sheet, and blob structures [5, 7, 11] in volume data.
- Analysis of filter responses for line structures using mathematical line models [7].
- Description and quantification (width and orientation measurement) of these local structures [10, 12].
- Analysis of sheet width quantification accuracy restricted by imaging resolution [17, 18].

For the multiscale enhancement, we design 3-D enhancement filters, which selectively respond to the specific type of local structures with specific size, based on the eigenvalues of the Hessian matrix of the Gaussian smoothed volume intensity function. The conditions that the eigenvalues need to satisfy for the local structures are analyzed to derive similarity measures to the local structures. We also design a multiscale integration scheme of the filter responses at different Gaussian SD values. The condition for the scale interval is analyzed to enhance equally over a specific range of structure sizes.

For the analysis of filter responses, mathematical line models with a noncircular cross section are used. The basic characteristics of the filter responses and their multiscale integration are analyzed simulationally. Although the line structure is considered here, the presented basic approach is applicable to filter responses to other local structures.

For the detection and quantification, we formulate a method using a second-order polynomial describing the local structures. We focus on line and sheet structures, which are basically characterized by their medial elements (medial surfaces and axes, respectively) and widths associated with the medial

elements. A second-order polynomial around a point of interest is defined by the gradient vector and Hessian matrix at the point. The medial elements are detected based on subvoxel localization of local maximum of the second-order polynomial within the voxel territory of a point of interest. The widths are measured along normal directions of the detected medical elements.

For the analysis of quantification accuracy, a theoretical approach is presented based on mathematical models of imaged local structures, imaging scanners, and quantification processes. Although the sheet structure imaged by MR scanners is focused here, the presented basic approach is applicable to accuracy analysis for different local structures imaged by either MR or CT scanners.

10.2 Multiscale Enhancement Filtering

10.2.1 Measures of Similarity to the Local Structures

Let $f(\vec{x})$ be an intensity function of a volume, where $\vec{x} = (x, y, z)$. Its Hessian matrix $\nabla^2 f$ is given by

$$\nabla^2 f(\vec{x}) = \begin{bmatrix} f_{xx}(\vec{x}) & f_{xy}(\vec{x}) & f_{xz}(\vec{x}) \\ f_{yx}(\vec{x}) & f_{yy}(\vec{x}) & f_{yz}(\vec{x}) \\ f_{zx}(\vec{x}) & f_{zy}(\vec{x}) & f_{zz}(\vec{x}) \end{bmatrix}, \tag{10.1}$$

where partial second derivatives of $f(\vec{x})$ are represented as $f_{xx}(\vec{x}) = \frac{\partial^2}{\partial x^2} f(\vec{x})$, $f_{yz}(\vec{x}) = \frac{\partial^2}{\partial y \partial z} f(\vec{x})$, and so on. The Hessian matrix $\nabla^2 f(\vec{x}_0)$ at $\vec{x}_0$ describes the second-order variations around x_0 [3–8, 11–14, 19]. The rotational invariant measures of the second-order local structure can be derived through the eigenvalue analysis of $\nabla^2 f(\vec{x}_0)$.

Let the eigenvalues of $\nabla^2 f(\vec{x})$ be $\lambda_1(\vec{x})$, $\lambda_2(\vec{x})$, $\lambda_3(\vec{x})$ ($\lambda_1(\vec{x}) \geq \lambda_2(\vec{x}) \geq \lambda_3(\vec{x})$), and their corresponding eigenvectors be $\vec{e}_1(\vec{x})$, $\vec{e}_2(\vec{x})$, $\vec{e}_3(\vec{x})$, respectively. The eigenvector $\vec{e}_1$, corresponding to the largest eigenvalue λ_1, represents the direction along which the second derivative is maximum, and λ_1 gives the maximum second-derivative value. Similarly, λ_3 and $\vec{e}_3$ give the minimum directional second-derivative value and its direction, and λ_2 and $\vec{e}_2$ the minimum directional second-derivative value orthogonal to $\vec{e}_3$ and its direction, respectively. λ_2 and $\vec{e}_2$ also give the maximum directional second-derivative value orthogonal to $\vec{e}_1$ and its direction.

λ_1, λ_2, and λ_3 are invariant under orthonormal transformations. λ_1, λ_2, and λ_3 are combined and associated with the intuitive measures of similarity to

Table 10.1: Basic conditions for each local structure and representative anatomical structures. Each structure is assumed to be brighter than the surrounding region

Structure	*Eigenvalue condition*	*Decomposed condition*	*Example(s)*
Sheet	$\lambda_3 \ll \lambda_2 \simeq \lambda_1 \simeq 0.$	$\lambda_3 \ll 0$	Cortex
		$\lambda_3 \ll \lambda_2 \simeq 0$	Cartilage
		$\lambda_3 \ll \lambda_1 \simeq 0$	
Line	$\lambda_3 \simeq \lambda_2 \ll \lambda_1 \simeq 0.$	$\lambda_3 \ll 0$	Vessel
		$\lambda_3 \simeq \lambda_2$	Bronchus
		$\lambda_2 \ll \lambda_1 \simeq 0$	
Blob	$\lambda_3 \simeq \lambda_2 \simeq \lambda_1 \ll 0.$	$\lambda_3 \ll 0$	Nodule
		$\lambda_3 \simeq \lambda_2$	
		$\lambda_2 \simeq \lambda_1$	

local structures. Three types of second-order local structures—sheet, line, and blob—can be classified using these eigenvalues. The basic conditions of these local structures and examples of anatomical structures that they represent are summarized in Table 10.1, which shows the conditions for the case where structures are bright in contrast with surrounding regions. Conditions can be similarly specified for the case where the contrast is reversed. Based on these conditions, measures of similarity to these local structures can be derived. With respect to the case of a line, we have already proposed a line filter that takes an original volume f into a volume of a line measure [7] given by

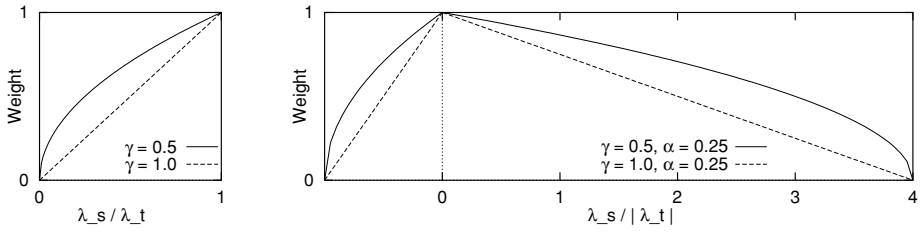
$$\mathcal{S}_{line}\{f\} = \begin{cases} |\lambda_3| \cdot \psi(\lambda_2;\lambda_3) \cdot \omega(\lambda_1;\lambda_2) & \lambda_3 \leq \lambda_2 < 0 \\ 0, & \text{otherwise,} \end{cases} \tag{10.2}$$

where ψ is a weight function written as

$$\psi(\lambda_s;\lambda_t) = \begin{cases} (\frac{\lambda_s}{\lambda_t})^{\gamma_{st}} & \lambda_t \leq \lambda_s < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{10.3}$$

in which γ_{st} controls the sharpness of selectivity for the conditions of each local structure (Fig. 10.1(a)), and ω is written as

$$\omega(\lambda_s;\lambda_t) = \begin{cases} (1 + \frac{\lambda_s}{|\lambda_t|})^{\gamma_{st}} & \lambda_t \leq \lambda_s \leq 0 \\ (1 - \alpha\frac{\lambda_s}{|\lambda_t|})^{\gamma_{st}} & \frac{|\lambda_t|}{\alpha} > \lambda_s > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{10.4}$$



(a) $\psi(\lambda_s;\ \lambda_t)$. (b) $\omega(\lambda_s;\ \lambda_t)$.

Figure 10.1: Weight functions in measures of similarity to local structures. (a) $\psi(\lambda_s;\lambda_t)$, representing the condition $\lambda_t \simeq \lambda_s$, where $\lambda_t \leq \lambda_s$. $\psi(\lambda_s;\lambda_t) = 1$ when $\lambda_t = \lambda_s$. $\psi(\lambda_s;\lambda_t) = 0$ when $\lambda_s = 0$. (b) $\omega(\lambda_s;\lambda_t)$, representing the condition $\lambda_t \ll \lambda_s \simeq 0$. $\omega(\lambda_s;\lambda_t) = 1$ when $\lambda_s = 0$. $\omega(\lambda_s;\lambda_t) = 0$ when $\lambda_t = \lambda_s \ll 0$ or $\lambda_s(\geq \frac{|\lambda_t|}{\alpha}) \gg 0$.

in which $0 < \alpha \leq 1$ (Fig. 10.1(b)). α is introduced in order to give $\omega(\lambda_s;\lambda_t)$ an asymmetrical characteristic in the negative and positive regions of λ_s.

Figure 10.2(a) shows the roles of weight functions in representing the basic conditions of the line case. In Eq. (10.2), $|\lambda_3|$ represents the condition $\lambda_3 \ll 0$, $\psi(\lambda_2;\lambda_3)$ represents the condition $\lambda_3 \simeq \lambda_2$ and decreases with deviation from the condition $\lambda_3 \simeq \lambda_2$, and $\omega(\lambda_1;\lambda_2)$ represents the condition $\lambda_2 \ll \lambda_1 \simeq 0$ and decreases with deviation from the condition $\lambda_1 \simeq 0$ which is normalized by λ_2. By multiplying $|\lambda_3|$, $\psi(\lambda_2;\lambda_3)$, and $\omega(\lambda_1;\lambda_2)$, we represent the condition for a line shown in Table 10.1. For the line case, the asymmetric characteristic of ω is based on the following observations:

- When λ_1 is negative, the local structure should be regarded as having a blob-like shape when $|\lambda_1|$ becomes large (lower right in Fig. 10.2(a)).
- When λ_1 is positive, the local structure should be regarded as being stenotic in shape (i.e., part of a vessel is narrowed), or it may be indicative of signal loss arising from the partial volume effect (lower left in Fig. 10.2(a)).

Therefore, when λ_1 is positive, we make the decrease with the deviation from the $\lambda_1 \simeq 0$ condition less sharp in order to still give a high response to a stenosis-like shape. We typically used $\alpha = 0.25$ and $\gamma_{st} = 0.5$ (or 1) in our experiments. Extensive analysis of the line measure, including the effects of parameters γ_{st} and α, can be found in [7].

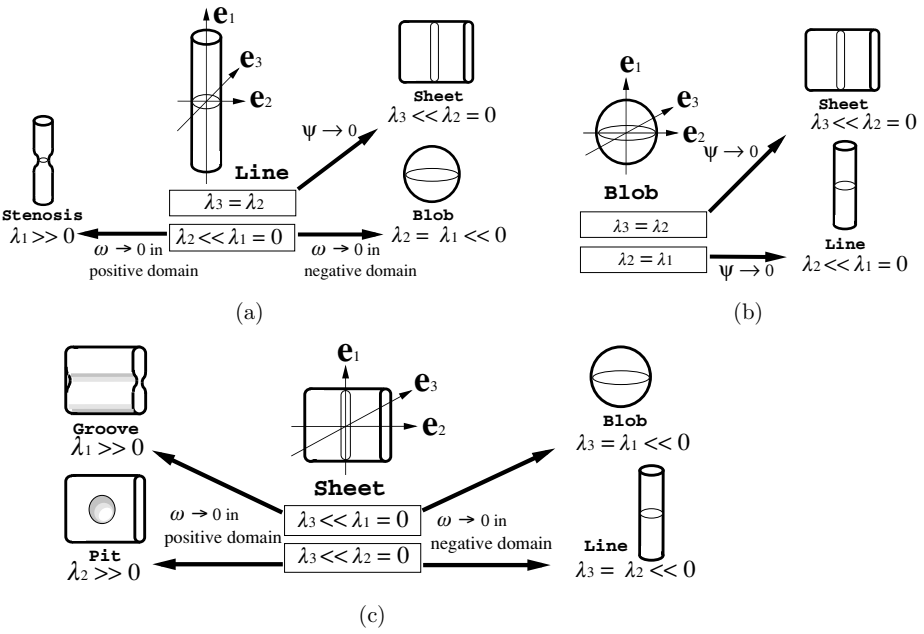


Figure 10.2: Schematic diagrams of measures of similarity to local structures. The roles of weight functions in representing the basic conditions of a local structure are shown. (a) Line measure. The structure becomes sheet-like and the weight function ψ approaches zero with deviation from the condition $\lambda_3 \simeq \lambda_2$, blob-like and the weight function ω approaches zero with transition from the condition $\lambda_2 \ll \lambda_1 \simeq 0$ to $\lambda_2 \simeq \lambda_1 \ll 0$, and stenosis-like and the weight function ω approaches zero with transition from the condition $\lambda_2 \ll \lambda_1 \simeq 0$ to $\lambda_1 \gg 0$. (b) Blob measure. The structure becomes sheet-like with deviation from the condition $\lambda_3 \simeq \lambda_2$, and line-like with deviation from the condition $\lambda_2 \simeq \lambda_1$. (c) Sheet measure. The structure becomes blob-like, groove-like, line-like, or pit-like with transition from $\lambda_3 \ll \lambda_1 \simeq 0$ to $\lambda_3 \simeq \lambda_1 \ll 0$, $\lambda_3 \ll \lambda_1 \simeq 0$ to $\lambda_1 \gg 0$, $\lambda_3 \ll \lambda_2 \simeq 0$ to $\lambda_3 \simeq \lambda_2 \ll 0$, or $\lambda_3 \ll \lambda_2 \simeq 0$ to $\lambda_2 \gg 0$, respectively. (© 2004 IEEE)

The specific shape given in Eq. (10.3) representing the condition $\lambda_t \simeq \lambda_s$ (where $t = 3$ and $s = 2$ for the line case) is based on the need to generalize the two line measures $\lambda_{min_{23}}$ and $\lambda_{g-mean_{23}}$ [3]:

$$\lambda_{min_{23}} = \begin{cases} \min(-\lambda_2, -\lambda_3) = -\lambda_2 & \lambda_2 < 0 \quad \text{and} \quad \lambda_3 < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{10.5}$$

and

$$\lambda_{g-mean_{23}} = \begin{cases} \sqrt{\lambda_2 \lambda_3} & \lambda_2 < 0 \quad \text{and} \quad \lambda_3 < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{10.6}$$

for the cases $\lambda_2 \leq 0$ and $\lambda_3 \leq 0$, $\lambda_{min_{23}}$ can be rewritten as

$$\lambda_{min_{23}} = -\lambda_2 = |\lambda_2| = |\lambda_3| \left(\frac{\lambda_2}{\lambda_3} \right), \tag{10.7}$$

and $\lambda_{g-mean_{23}}$ as

$$\lambda_{g-mean_{23}} = \sqrt{\lambda_2 \lambda_3} = |\lambda_3| \left(\frac{\lambda_2}{\lambda_3} \right)^{0.5}. \tag{10.8}$$

These measures take into account the conditions $\lambda_3 \ll 0$ and $\lambda_3 \simeq \lambda_2$. $|\lambda_3| \cdot \psi(\lambda_2; \lambda_3)$ in Eq. (10.2) is equal to $\sqrt{\lambda_3 \lambda_2}$ and $-\lambda_2$ when $\gamma_{23} = 0.5$ and $\gamma_{23} = 1$, respectively. In this formulation [7], the same type of function shape as that in Eq. (10.3) is used for Eq. (10.4) to add the condition $\lambda_2 \ll \lambda_1 \simeq 0$.

We can extend the line measure to the blob and sheet cases. In the blob case, the condition $\lambda_3 \simeq \lambda_2 \simeq \lambda_1 \ll 0$ can be decomposed into $\lambda_3 \ll 0$ and $\lambda_3 \simeq \lambda_2$ and $\lambda_2 \simeq \lambda_1$. By representing the condition $\lambda_t \simeq \lambda_s$ using $\psi(\lambda_s; \lambda_t)$, we can derive a blob filter given by

$$\mathcal{S}_{blob}\{f\} = \begin{cases} |\lambda_3| \cdot \psi(\lambda_2; \lambda_3) \cdot \psi(\lambda_1; \lambda_2) & \lambda_3 \leq \lambda_2 \leq \lambda_1 < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{10.9}$$

In the sheet case, the condition $\lambda_3 \ll \lambda_2 \simeq \lambda_1 \simeq 0$ can be decomposed into $\lambda_3 \ll 0$ and $\lambda_3 \ll \lambda_2 \simeq 0$ and $\lambda_3 \ll \lambda_1 \simeq 0$. By representing the condition $\lambda_t \ll \lambda_s \simeq 0$ using $\omega(\lambda_s; \lambda_t)$, we can derive a sheet filter given by

$$\mathcal{S}_{sheet}\{f\} = \begin{cases} |\lambda_3| \cdot \omega(\lambda_2; \lambda_3) \cdot \omega(\lambda_1; \lambda_3) & \lambda_3 < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{10.10}$$

Figures 10.2(b) and 10.2(c) show the relationships between the eigenvalue conditions and weight functions in the blob and sheet measures.

10.2.2 Multiscale Computation and Integration of Filter Responses

Local structures can exist at various scales. For example, vessels and bone cortices can, respectively, be regarded as line and sheet structures with various widths. In order to make filter responses tunable to a width of interest,

the derivative computation for the gradient vector and the Hessian matrix is combined with Gaussian convolution. By adjusting the standard deviation of Gaussian convolution, local structures with a specific range of widths can be enhanced. The Gaussian function is known as a unique distribution optimizing localization in both the spatial and frequency domains [20]. Thus, convolution operations can be applied within local support (due to spatial localization) with minimum aliasing errors (due to frequency localization).

We denote the local structure filtering for a volume blurred by Gaussian convolution with a standard deviation σ_f as

$$\mathcal{S}_\xi\{f;\sigma_f\}, \tag{10.11}$$

where $\xi \in \{sheet, line, blob\}$. The filter responses decrease as σ_f in the Gaussian convolution increases unless appropriate normalization is performed [21–23]. In order to determine the normalization factor, we consider a Gaussian-shaped model of sheet, line, and blob with variable scales.

Sheet, line, and blob structures with variable widths are modeled as

$$\ell_{sheet}(\vec{x};\sigma_r) = \exp\left(-\frac{x^2}{2\sigma_r^2}\right), \tag{10.12}$$

$$\ell_{line}(\vec{x};\sigma_r) = \exp\left(-\frac{x^2+y^2}{2\sigma_r^2}\right), \tag{10.13}$$

and

$$\ell_{blob}(\vec{x};\sigma_r) = \exp\left(-\frac{x^2+y^2+z^2}{2\sigma_r^2}\right), \tag{10.14}$$

respectively, where σ_r controls the width of the structures.

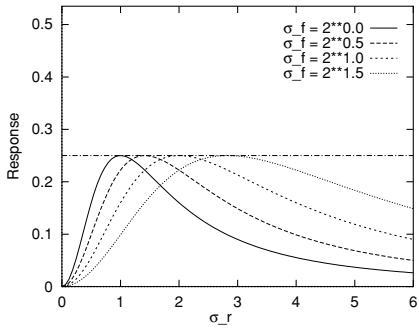
We determine the normalization factor so that $\mathcal{S}_\xi\{\ell_\xi(\vec{x};\sigma_r);\sigma_f\}$ satisfies the following condition:

- $\max_{\sigma_r} \mathcal{S}_\xi\{\ell_\xi(\vec{0};\sigma_r);\sigma_f\}$ is constant, irrespective of σ_f, where $\vec{0} = (0,0,0)$.

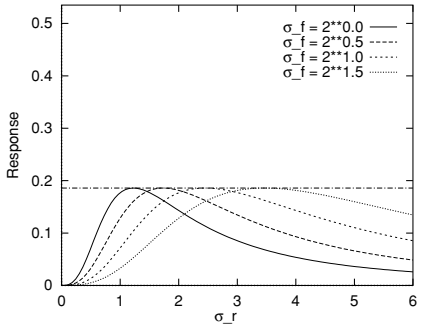
The above condition can be satisfied when the Gaussian second derivatives are computed by multiplying by σ_f^2 as the normalization factor. That is, the normalized Gaussian derivatives are given by

$$f_{x^p y^q z^r}(\vec{x};\sigma_f) = \left\{\sigma_f^2 \cdot \frac{\partial^2}{\partial x^p \partial y^q \partial z^r}\mathrm{Gauss}(\vec{x};\sigma_f)\right\} * f(\vec{x}) \tag{10.15}$$

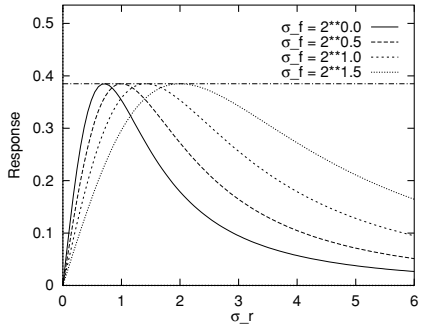
where p, q, and r are non-negative integer values satisfying $p+q+r=2$, and $\mathrm{Gauss}(\vec{x};\sigma)$ is an isotropic 3D Gaussian function with a standard deviation σ given by $(\sqrt{2\pi}\sigma)^{-1}\exp(-|\vec{x}|^2/(2\sigma^2))$ (see the Questions section at the end of



(a) Line



(b) Blob.



(c) Sheet.

Figure 10.3: Plots of normalized responses of local structure filters for corresponding local models, $\mathcal{S}_\xi\{\ell_\xi(\vec{0};\sigma_r);\sigma_f\}$, where σ_r is continuously varied, and $\sigma_f = \sigma_i s^{i-1}$ ($\sigma_1 = 1, s = 1.414$, and $i = 1, 2, 3, 4$). See "Brain Storming Questions" at the end of this chapter for the theoretical derivations of the response curves. (a) Response of the line filter for the line model ($\xi = line$). (b) Response of the blob filter for the blob model ($\xi = blob$). (c) Response of the sheet filter for the sheet model ($\xi = sheet$). (© 2004 IEEE)

this chapter for the derivation). Figure 10.3 shows the normalized response of $\mathcal{S}_\xi\{\ell_\xi(\vec{0};\sigma_r);\sigma_f\}$ (where $\sigma_f = \sigma_i s^{i-1}$, $\sigma_1 = 1$, $s = \sqrt{2}$, and $i = 1, 2, 3, 4$) for $\xi \in \{sheet, line, blob\}$ when σ_r is varied.

In the line case, the maximum of the normalized response $\mathcal{S}_{line}\{\ell_{line}(\vec{0},\sigma_r);\sigma_f\}$ is $\frac{1}{4}(= 0.25)$ when $\sigma_r = \sigma_f$ [7]. That is, $\mathcal{S}_{line}\{f;\sigma_f\}$ is regarded as being tuned to line structures with a width $\sigma_r = \sigma_f$. A line filter with a single scale gives a high response in only a narrow range of widths. We call the curves shown in Fig. 10.3 as *width response curves*, which represent filter characteristics like frequency response curves. The width response curve

of the line filter can be adjusted and widened using multiscale integration of filter responses given by

$$\mathcal{M}_{line}\{f;\sigma_1,s,n\} = \max_{1\le i\le n} \mathcal{S}_{line}\{f;\sigma_i\}, \tag{10.16}$$

where $\sigma_i = s^{i-1}\sigma_1$, in which σ_1 is the smallest scale, s is a scale factor, and n is the number of scales [7]. The width response curve of multiscale integration using the four scales consists of the maximum values among the four single-scale width response curves, and gives nearly uniform responses in the width range between $\sigma_r = \sigma_1$ and $\sigma_r = \sigma_4$ when $s = \sqrt{2}$ (Fig. 10.3(a)). While the width response curve can be perfectly uniform if continuous variation values are used for σ_f, the deviation from the continuous case is less than 3% using discrete values for σ_f with $s = \sqrt{2}$ [7]. Similarly, in the cases of $\mathcal{S}_{sheet}\{\ell_{sheet}(\vec{0},\sigma_r);\sigma_f\}$ and $\mathcal{S}_{blob}\{\ell_{blob}(\vec{0},\sigma_r);\sigma_f\}$, the maximum of the normalized response is $\frac{2}{(\sqrt{3})^3}(\approx 0.385)$ when $\sigma_r = \frac{\sigma_f}{\sqrt{2}}$ (Fig. 10.3(b)), and $\frac{2}{3}(\sqrt{\frac{3}{5}})^5(\approx 0.186)$ when $\sigma_r = \sqrt{\frac{3}{2}}\sigma_f$ (Fig. 10.3(c)), respectively (see the Question section at the end of this chapter for the derivation). For the second-order cases, the width response curve can be adjusted and widened using the multiscale integration method given by

$$\mathcal{M}_{\xi}\{f;\sigma_1,s,n\} = \max_{1\le i\le n} \mathcal{S}_{\xi_2}\{f;\sigma_i\}, \tag{10.17}$$

where $\xi \in \{sheet, line, blob\}$.

10.2.3 Implementation Issues

10.2.3.1 Sinc Interpolation Without Gibbs Ringing

Our 3-D local structure filtering methods described above assume that volume data with isotropic voxels are used as input data. However, voxels in medical volume data are usually anisotropic since they generally have lower resolution along the third direction, i.e., the direction orthogonal to the slice plane, than within slices. Rotational invariant feature extraction becomes more intuitive in a space where the sample distances are uniform. That is, structures of a particular size can be detected on the same scale independent of the direction when the signal sampling is isotropic. We therefore introduce a preprocessing procedure for 3-D local structure filtering in which we perform interpolation to make each voxel isotropic. Linear and spline-based interpolation methods are often used, but blurring is inherently involved in these approaches. Because, as noted above,

the original volume data is inherently blurrier in the third direction, further degradation of the data in that direction should be avoided. For this reason, we opted to employ sinc interpolation so as not to introduce any additional blurring. After Gaussian-shaped slopes are added at the beginning and end of each profile in the third direction to avoid unwanted Gibbs ringing, sinc interpolation is performed by zero-filled expansion in the frequency domain [24, 25].

The method for sinc interpolation without Gibbs ringing is described below. The sinc interpolation along the third (z-axis) direction is performed by zero-filled expansion in the frequency domain. Let $f(i)$ ($i = 0, 1, \ldots, n-1$) be the profile in the third direction. In the discrete Fourier transform of $f(i)$, $f(i)$ should be regarded as cyclic and then $f(n-1)$ and $f(0)$ are essentially adjacent. Unwanted Gibbs ringing occurs in the interpolated profile due to the discontinuity between $f(n-1)$ and $f(0)$. Thus, Gaussian-shaped slopes are added at the beginning and end of $f(i)$ to avoid the occurrence of unwanted ringing before the sinc interpolation. Let $f'(i)$ ($i = -3 \cdot \sigma, \ldots, 0, 1, \ldots, n-1, n, \ldots, 3 \cdot \sigma + n$) be the modified profile, which is given by

$$f'(i) = \begin{cases} \exp(-\frac{i^2}{2\sigma^2}) \cdot f(0) & i = -3 \cdot \sigma, \ldots, 0 \\ f(i) & i = 0, \ldots, n-1 \\ \exp(-\frac{(i-n+1)^2}{2\sigma^2}) \cdot f(n-1) & i = n, \ldots, 3 \cdot \sigma \end{cases} \tag{10.18}$$

where the variation is sufficiently smooth everywhere, including between $f(3 \cdot \sigma + n)$ and $f(-3 \cdot \sigma)$. The discrete Fourier transform of $f'(i)$ is performed (we used $\sigma = 4$). After the sinc interpolation of $f'(i)$, the added Gaussian-shaped slopes are removed.

10.2.3.2 Computation of Gaussian Derivatives and Eigenvalues

The computation of the Gaussian derivatives in the Hessian matrix and the gradient vector (needed in the later chapters) can be implemented using three separate convolutions with 1-D kernels as represented by

$$\begin{aligned} f_{x^p y^q z^r}(\vec{x}; \sigma_f) &= \left\{ \frac{\partial^2}{\partial x^p \partial y^q \partial z^r} \mathrm{Gauss}(\vec{x}; \sigma_f) \right\} * f(\vec{x}) \\ &= \frac{\mathrm{d}^p}{\mathrm{d}x^p} \mathrm{Gauss}(x; \sigma_f) * \left\{ \frac{\mathrm{d}^q}{\mathrm{d}y^q} \mathrm{Gauss}(y; \sigma_f) \right. \\ &\quad \left. * \left\{ \frac{\mathrm{d}^r}{\mathrm{d}z^r} \mathrm{Gauss}(z; \sigma_f) * f(\vec{x}) \right\} \right\} \end{aligned} \tag{10.19}$$

where p, q, and r are nonnegative integers satisfying $p+q+r \leq 2$. To obtain the normalized Gaussian derivatives, σ_f^{p+q+r} further needs to be multiplied with $f_{x^p y^q z^r}(\vec{x}\,;\sigma_f)$. In our experience, it is recommended that the radius of the kernel should be $3 \cdot \sigma_f$, $4 \cdot \sigma_f$, and $5 \cdot \sigma_f$ in simple smoothing ($p+q+r=0$), first ($p+q+r=1$), and second derivatives ($p+q+r=2$), respectively, for the accurate computation of the Gaussian derivatives and smoothing. Using this decomposition, the amount of computation needed can be reduced from $O(n^3)$ to $O(3n)$, where n is the kernel diameter.

The eigenvalues were computed using Jacobi's method in the simulations and experiments shown in this chapter. A numerical problem could exceptionally occur in this computation for synthesized images of the mathematical line models without noise. However, the numerical problem can be avoided by adding a very small Gaussian noise. We did not have such a numerical problem in the experiments using MR and CT images since noise is essentially involved in real images.

10.2.3.3 Guidelines for Parameter Value Selection

The multiscale enhancement filter includes several parameters. In defining the similarity measures to local structures, γ_{st} and α in Eqs. (10.3) and (10.4) need to be specified, while in the multiscale integration, the smallest scale σ_1, the scale factor s, and the number of scale levels n in Eq. (10.17) need to be specified.

With regard to the parameters in the line measure, if there are strong sheet structures to be removed from an image, γ_{23} should be 1.0. If the cross sections of line structures of interest vary not only in size but also in shape, γ_{23} should be 0.5. γ_{12} also should be determined to optimize the trade-off between the preservation of branches and the reduction of noise and spurious branches. However, the performance has been found to be relatively insensitive to the values of γ_{23} and γ_{12} as long as they are between 0.5 and 1.0. Also, we have empirically found that $\alpha = 0.25$ can be considered a good compromise.

With respect to the parameters in multiscale integration, we experimentally found that s should be 1.5, $\sqrt{2}$, or less for reasonable approximation of the multiscale integration of the continuous scale. The minimum value of σ_1 for discrete samples of volume data was 0.8 (voxels), which are experimentally shown in [7]. Considering the above observations, suitable values for σ_1 and n should be found out for each case taking account of the anatomical structure of

interest based on the width response curves as shown in Fig. 10.3. We confirmed that the results were quite stable for different images obtained under similar conditions once suitable values have been determined.

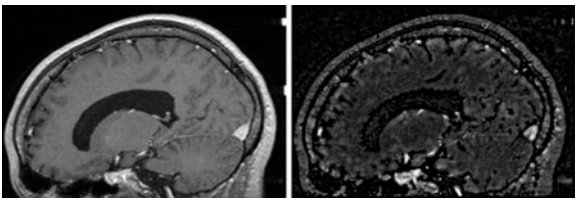
The detailed analyses of the effects of parameter values on the filter responses, which are the bases of the above guidelines, are discussed for the line case in the next section, and more thorough analyses of them are found in [7].

10.2.4 Examples

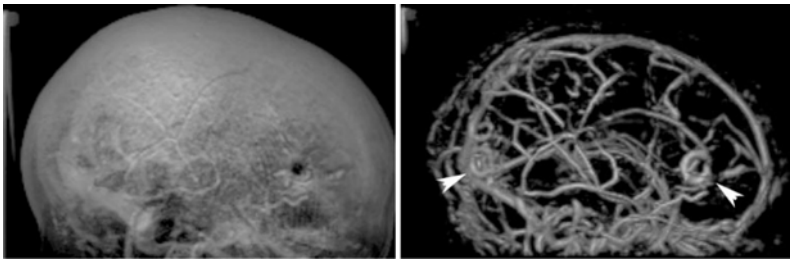
10.2.4.1 Neurovascular Visualization from 3-D MR Data

Multiscale line filtering was applied to postcontrast gradient-echo (SPGR) MR images of the brain for the purpose of vessel enhancement so as to use the enhanced vein structures as landmarks for brain tumor resection [26,27]. The MRI dataset consisted of 192 sagittal slices of 256×256 pixels. The pixel dimensions were 1.0 mm^2, while the slice thickness was 1.2 mm. The DSA images obtained from the same patient were also available, and used to compare with the vessel visualization with the MRI data. We used 80 slices corresponding to the right half of the head, and trimmed a region of 220×150 pixels from each slice. Sinc interpolation was performed for the trimmed images to obtain isotropic voxel sampling (the left frame of Fig. 10.4(a)). Multiscale line filtering was applied to the interpolated images using $\gamma_{23} = \gamma_{12} = 1.0$, $\alpha = 0.25$, $\sigma_1 = 0.8$ pixel, $s = 1.5$, and $n = 3$ (the right frame of Fig. 10.4(a)). The original and line-filtered images were visualized using the volume-rendering technique [28,29] (Fig. 10.4(b)).

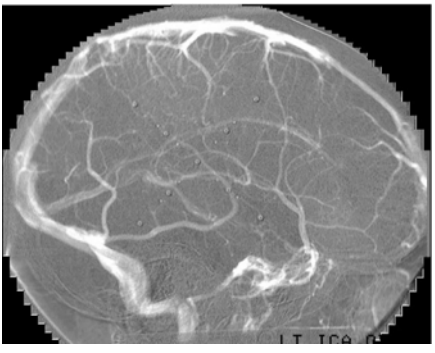
It is quite difficult to perceive the vessels through the skin in the volume-rendered original images, but almost the same vein structures as in the DSA image (Fig. 10.4(c)) can be clearly seen in the volume-rendered line-filtered images. Although the vessels and the skin had almost the same intensity level in the original images, the line filter could mostly remove the effect of the skin since it has a sheet-like structure. As a side effect, the line filter also gave high levels of response to the rims of biopsy holes in the skin and skull. It is a current limitation of our formulation that the line filter gives a high level of response to rim structures as well as to line structures. In order to remove rim structures, a procedure such as the nonlinear combination of the first derivatives for line detection from 2-D images [3, 30], needs to be extended for use with 3-D images.



(a)



(b)

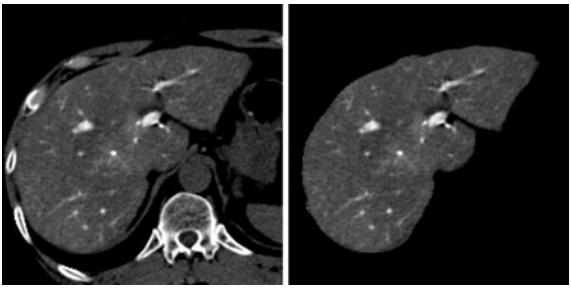


(c)

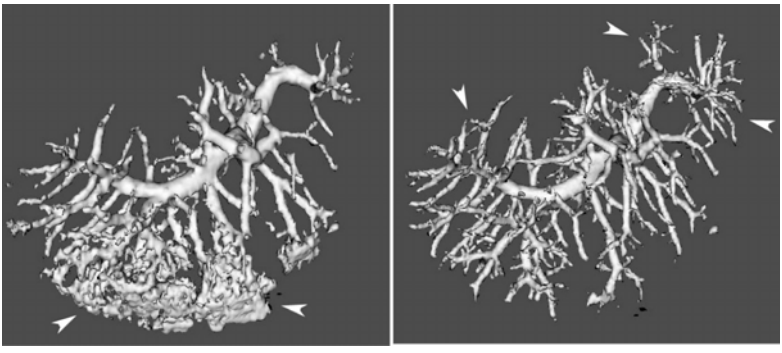
Figure 10.4: Neurovascular visualization from 3-D MR data. (a) Original (left) and line-filtered (right) cross-sectional images. (b) Original (left) and line-filtered (right) volume rendered images. (c) DSA (digital subtraction angiography) image at a vein phase.

10.2.4.2 Portal Vein Segmentation from 3-D CT Data

Multiscale line filtering was applied to abdominal CT images taken by a helical CT scanner so as to segment the portal veins to localize a tumor with the relation to them for surgical planning. The CT dataset consisted of 43 slices of 512 $\times$ 512 pixels; the pixel dimensions were 0.59 mm^2. The beam width was 3 mm and the reconstruction pitch was 2.5 mm. The CT data were imaged using CTAP



(a)



(b)

Figure 10.5: Liver vessel (portal vein) segmentation from abdominal CT images. (a) Original cross-sectional images (left) and segmented liver region (right). (b) Original (left) and line-filtered (right) surface-rendered images.

(CT arterial portography)[2]; the portal veins had high CT values due to the injection of contrast material. A region of 400 $\times$ 400 pixels from each slice was trimmed, which included the whole liver (the left frame of Fig. 10.5(a)), and further the image size was reduced to half using the Laplacian pyramid [31] to reduce a computational amount to a practical level. The liver regions were roughly hand-segmented by a radiology specialist and used as a mask (the right frame of Fig. 10.5(a)). The CT values were converted so that the image intensity f was zero for less than $f_{\min}$, $f_{\max} - f_{\min}$ for more than $f_{\max}$, and $f - f_{\min}$ for between $f_{\min}$ and $f_{\max}$ (where $f_{\min} = 1000$ and $f_{\max} = 1300$). Line filtering was applied to the sinc-interpolated images using $\gamma_{23} = \gamma_{12} = 1.0$, $\alpha = 0.25$, $\sigma_1 = 0.8$ pixels,

[2]The CT data were obtained by a helical CT scanner with the 20-sec delay following the administration of contrast material using a catheter inserted in the SMA (superior mesenteric artery). This method of portal vein imaging is called CTAP (CT arterial portography).

$s = 1.5$, and $n = 2$. We multiplied the mask images with the line-filtered images, thresholded the masked line-filtered images using an appropriate threshold value, and removed small connected components whose size was less than 10 voxels.

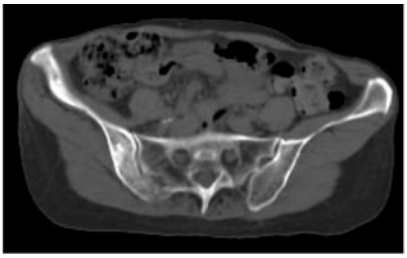
In Fig. 10.5(b), the left frame gives the rendered result of the original binary images, and the right frame shows a combination of the line-filtered binary images for small-vessel detection and the original binary images using relatively high threshold values for large-vessel detection. The two binary images were combined by taking the union of them. The CT data were scanned when the contrast material in the portal vein began to be absorbed by the liver tissues, as seen in the lower part of Fig. 10.5(b). Such a condition is quite common in CTAP for portal vein imaging. In the original images, the small vessels appear buried due to the contrast material absorbed by the liver tissue. In the combined result of the original and line-filtered images, not only is the nonuniformity of the contrast material canceled out, but also the recovery of small vessels is significantly improved over the entire liver area.

10.2.4.3 Pelvic Bone Tumor and Cortex Visualization from 3-D CT Data

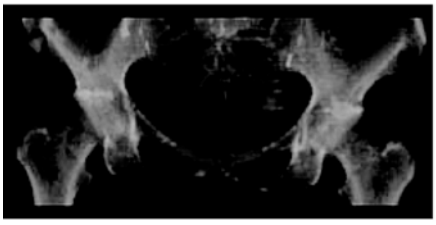
A single-scale sheet filter was applied to pelvic CT images for bone cortex enhancement. The purpose was to visualize the distribution of bone tumors and localize them in relation to the pelvic structure for biopsy planning as well as diagnosis [32]. Healthy bone cortex tissues and bone tumors have similar original CT values. However, bone cortices are sheet-like in structure, while tumors are not. Thus, enhanced bone cortices using sheet enhancement filtering are expected to be discriminated from tumors which are not enhanced.

The CT dataset consisted of 40 slices with a 512×512 matrix (Fig. 10.6(a)); the pixel dimensions were 0.82 mm^2. The slice thickness and reconstruction pitch were 5 mm. The matrix was reduced to half in the xy-plane, and thus the pixel interval was 1.64 mm. Sheet filtering was applied to the sinc-interpolated images using $\sigma_f = 1.0$ pixel, $\gamma_{23} = \gamma_{13} = 0.5$, and $\alpha = 0.25$.

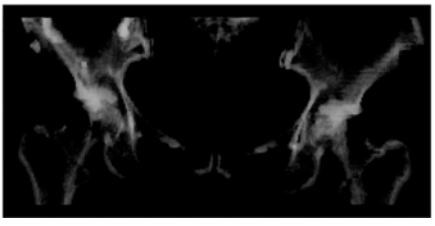
Figure 10.6(b) shows the color volume renderings of bone tumors (pink) and cortices (white). In the left frame, the opacity and color functions were adjusted using only CT values of the original images. In the right frame, both the original and sheet filtered images were used, where voxels having high intensities



(a)



(b)



(c)

Figure 10.6: Visualization of pelvic bone tumors from CT data. (a) Original CT slice image. (b) Volume rendered images of bone tumors and cortices. Left: Using only original images. Right: Using original and sheet-filtered images. (c) Manually traced tumor regions. (© 2004 IEEE). A color version of this figure will appear on the CD that accompanies the volume.

both in sheet-filtered and original images were assigned as cortices (white), and those having high in original but low in sheet-filtered images were assigned as tumors (pink). Figure 10.6(c) shows the rendered color image generated from the tumor regions manually traced by a radiology specialist, which is regarded as an ideal visualization. The color rendering of the left frame of Fig. 10.6(b) was well correlated with Fig. 10.6(c) (the "ideal" image), and the bone tumors were visualized considerably better than only using original CT images. However, nontumor regions around articular spaces were also detected mainly due to the partial volume effect (by a large slice thickness).

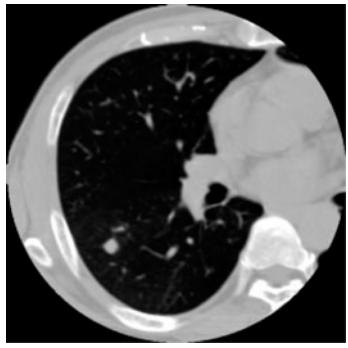
10.2.4.4 Lung Nodule and Vessel Visualization from 3-D CT Data

Multiscale blob and line filters were applied to chest CT images for nodule enhancement and vessel enhancement to detect early-stage lung cancers and visualize them with relation to peripheral vessels [33–35]. The CT dataset consisted of 60 slices with a 512×512 matrix (Fig. 10.7(a)); the pixel dimensions were 0.39 mm^2. The slice thickness and reconstruction pitch were 2 mm and 1 mm, respectively. The matrix was reduced to half in the xy-plane, and thus the pixel interval was 0.78 mm. The data were then interpolated along the z-axis using sinc interpolation so that the voxel was isotropic. While nodules, vessels, and other soft tissues have similar CT values in original images, the nodules and vessels have blob and line structures, respectively. Multiscale blob filtering was applied to the interpolated images using $\gamma_{23} = \gamma_{12} = 0.5$, $\alpha = 0.25$, $\sigma_1 = 2.0$ pixels, $s = \sqrt{2}$, and $n = 3$. Multiscale line filtering was applied using $\gamma_{23} = \gamma_{12} = 1.0$, $\alpha = 0.25$, $\sigma_1 = 1.0$ pixels, $s = \sqrt{2}$, and $n = 3$.

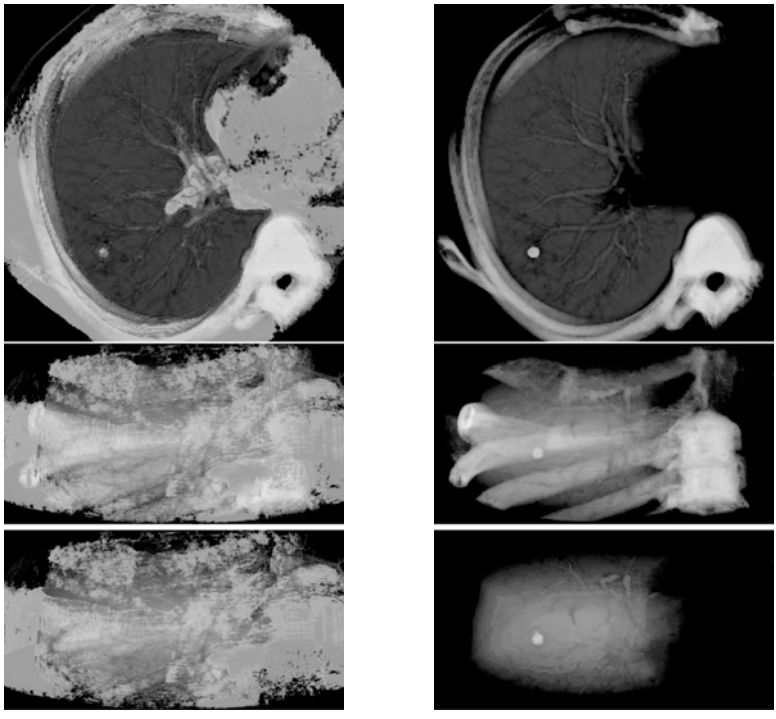
Figure 10(b) shows the color volume renderings of nodules (green), vessels (red), lung (violet), and bone tissues (white). In the left frame, the opacity and color functions were adjusted using only CT values of the original images. In the right frame, the original, blob, and line filtered images were used, where voxels having high intensities in the blob-filtered images were assigned as nodules (green), those having high in the line-filtered images as vessels (red), those having low in the original and two filtered images as lung tissues (violet), and those having high in the original but low in the two filtered images as bone tissues (white). The nodules and vessels were clearly depicted with different colors using blob and line enhancement filtering, while it was difficult to discriminate soft tissues into different categories using only original intensity values.

10.3 Analysis of Multiscale Line Filter Responses

The measures of similarity to the local structures have been introduced based on the ideal local structures with an isotropic Gaussian cross section shown in Eqs. (10.12)–(10.14). To examine the effects of parameters involved in the



(a)



(b)

Figure 10.7: Visualization of lung nodule and vessel from CT data (Color Slide). (a) Original CT slice image. (b) Volume rendered images of nodules (green), vessels (red), lung (violet), and bone (white) tissues. Left: Using only original images. Right: Using original, blob-filtered, and line-filtered images. (© 2004 IEEE).

multiscale enhancement filters, simulation experiments were performed using synthesized images. In this section, the effects of parameter γ_{23}, initial scale σ_0, and scale factor s are presented based on simulations using a line model with an elliptic cross section. The effects of parameters γ_{12} and α are not presented here. Simulation experiments are presented in [7] to analyze the effects of parameters γ_{12}, and α using curved and branched line models.

10.3.1 Single-Scale Filter Responses to Mathematical Line Models

The line measure generalizes $\lambda_{min_{23}}$ in Eq. (10.5) and $\lambda_{g-mean_{23}}$ in Eq. (10.6). An alternative measure is to use the arithmetic mean of $-\lambda_2$ and $-\lambda_3$, which is given by

$$\lambda_{a-mean_{23}} = -\frac{\lambda_2 + \lambda_3}{2}. \tag{10.20}$$

To compare these three measures, let us consider a 3-D line image with elliptic (nonisotropic Gaussian) cross sections given by

$$\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y) = \exp\left\{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right\}. \tag{10.21}$$

When $\sigma_x = \sigma_y$, $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$ can be regarded as an ideal line, that is, $\ell_{line}(\vec{x}; \sigma_x)$. Figure 10.8 shows the plots of the three measures and the eigenvalue variations of the Hessian matrix along the x-axis for the ideal line ($\sigma_x = \sigma_y = 4$ in Eq. (10.21), $\sigma_f = 4$) and the sheet-like ($\sigma_x = 20$, $\sigma_y = 3$, $\sigma_f = 4$) cases. The directions of the three eigenvectors at the points on the x-axis are identical to the x-axis, y-axis, and z-axis in the 3-D images modeled by Eq. (10.21). Let $\vec{e}_x$, $\vec{e}_y$, and $\vec{e}_z$ be eigenvectors whose directions are identical to the x-axis, y-axis, and z-axis, respectively, and let λ_x, λ_y, and λ_z be the respective corresponding eigenvalues. Figures 10.8(a) and 10.8(b) show the plots of $\lambda_{g-mean_{23}}$, $\lambda_{a-mean_{23}}$, and $\lambda_{min_{23}}$ as well as the original profiles, while Figs. 10.8(c) and 10.8(d) show the plots of λ_x, λ_y, and λ_z. In the ideal line case, both λ_2 and λ_3 are negative with large absolute values near the line centers. Since λ_3 tends to have a larger absolute value in the sheet-like case than in the line case (Fig. 10.8(d)), $\lambda_{a-mean_{23}}$ still gives a high response in the sheet-like case even if λ_2 has a small absolute value. Figure 10.9 shows the responses of the three measures at the center of

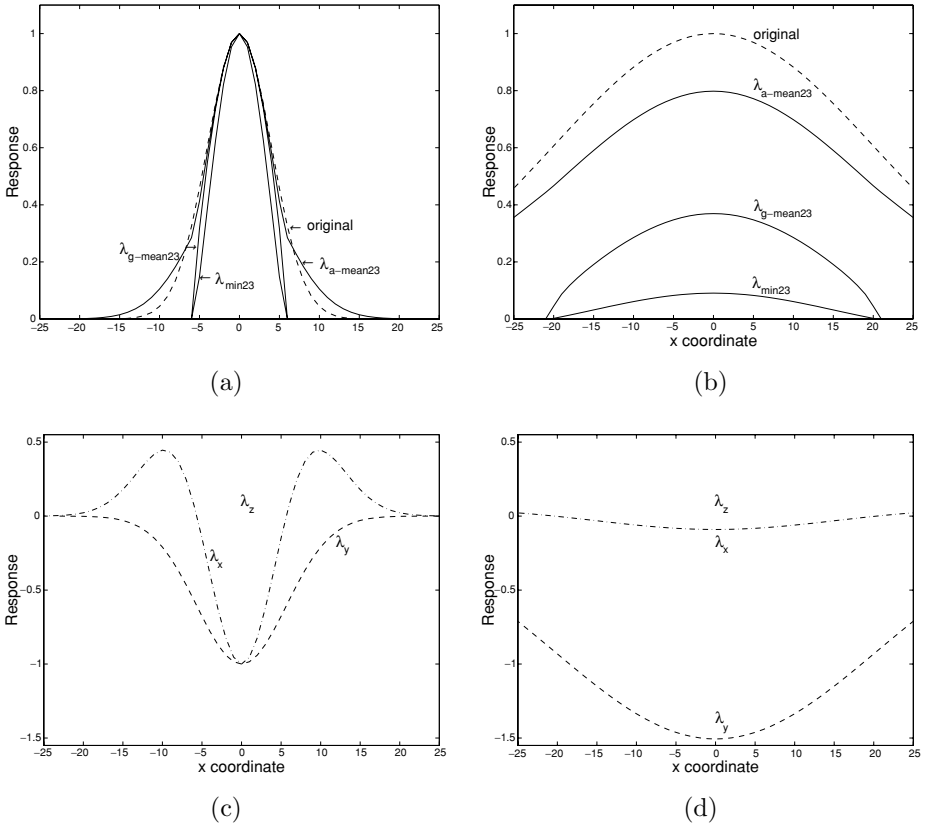
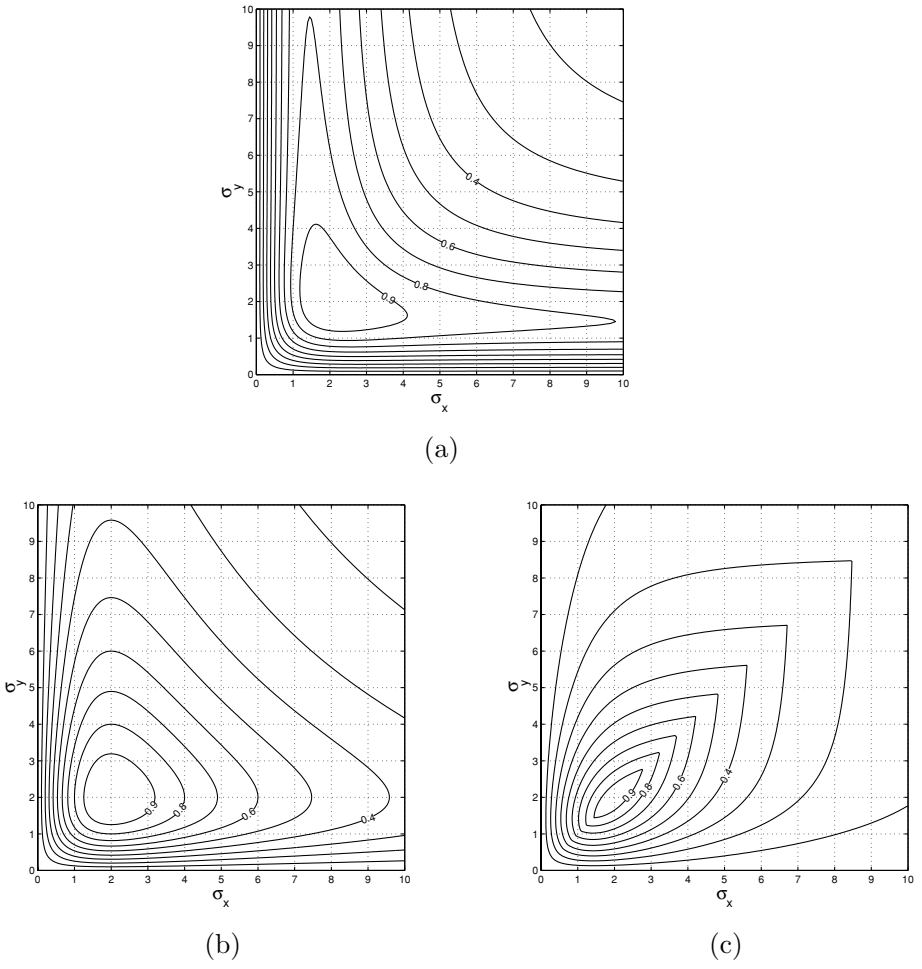


Figure 10.8: Responses of eigenvalues and 3-D line filters to $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$ along the x-axis. The eigenvalues and filter responses are normalized so that $|\lambda_2|$ and $|\lambda_3|$ are one at $x = y = 0$ when $\sigma_x = \sigma_y = \sigma_f = 4$. (a) $\lambda_{g-mean_{23}}$, $\lambda_{a-mean_{23}}$, $\lambda_{min_{23}}$, and the original profile for the ideal line case ($\sigma_x = \sigma_y = 4$ in $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$, $\sigma_f = 4$). (b) $\lambda_{g-mean_{23}}$, $\lambda_{a-mean_{23}}$, $\lambda_{min_{23}}$, and the original profile for the sheet-like case ($\sigma_x = 20$ and $\sigma_y = 3$ in $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$, $\sigma_f = 4$). (c) Eigenvalues for the ideal line case. (d) Eigenvalues for the sheet-like case.

the line ($x = y = 0$) when σ_x and σ_y in Eq. (10.21) are varied. While $\lambda_{min_{23}}$ and $\lambda_{g-mean_{23}}$ decrease with deviations from the conditions $\sigma_x \approx \sigma_f$ and $\sigma_y \approx \sigma_f$, $\lambda_{a-mean_{23}}$ gives high responses if $\sigma_x \approx \frac{\sigma_f}{\sqrt{2}}$ or $\sigma_y \approx \frac{\sigma_f}{\sqrt{2}}$. Thus, $\lambda_{a-mean_{23}}$ gives relatively high responses to sheet-like structures while $\lambda_{min_{23}}$ ($\gamma_{23} = 1$ in Eq. (10.3)) and $\lambda_{g-mean_{23}}$ ($\gamma_{23} = 0.5$ in Eq. (10.3)) are able to discriminate line structures from sheet-like structures.



(a)

(b) (c)

Figure 10.9: Responses of 3-D line filters to $\ell_{elliptic}(\vec{x}\,;\sigma_x,\sigma_y)$ at $x = y = 0$ with variable σ_x and σ_y at the center of the line. The responses are normalized so that $|\lambda_2|$ and $|\lambda_3|$ are one at $x = y = 0$ when $\sigma_x = \sigma_y = \sigma_f = 2$. (a) $\lambda_{a-mean_{23}}$. (b) $\lambda_{g-mean_{23}}$. (c) $\lambda_{min_{23}}$.

10.3.2 Multiscale Responses of Continuous Scale Integration

In order to analyze multiscale filter responses and develop design criteria for the multiscale integration, we observe the variations in the height and width of the original line image $\ell_{line}(\vec{x}\,;\sigma_r)$ in Eq. (10.13) and its filter responses. The

multiscale response for $\ell_{line}(\vec{x};\sigma_r)$ is given by

$$\mathcal{M}_{line}\{\ell_{line}(\vec{x};\sigma_r)\} = \max_{\sigma_f} \mathcal{S}_{line}\{\ell_{line}(\vec{x};\sigma_r);\sigma_f\}, \tag{10.22}$$

We define the height measure of the multiscale filter response as the peak response $h_{\mathrm{M}}(\sigma_r) = \mathcal{M}_{line}\{\ell_{line}(0,0,z;\sigma_r)\}$. Since the filter response is normalized, $h_{\mathrm{M}}(\sigma_r)$ is constant regardless of σ_r. That is,

$$h_{\mathrm{M}}(\sigma_r) = h_{\mathrm{M}_c}, \tag{10.23}$$

where $h_{\mathrm{M}_c} = 0.25$ (see the "Brain Storming Question" at the end of this chapter for the derivation). We define the width measure $w_{\mathrm{M}}(\sigma_r)$ of the multiscale filter response as the distance $\sqrt{x_0^2 + y_0^2}$ from the z-axis to the circular locus where $\mathcal{M}_{line}\{\ell_{line}(x_0, y_0, z;\sigma_r)\}$ gives half of the peak response, that is, $\frac{h_{\mathrm{M}_c}}{2}$. Let $w'_{\mathrm{M}}(\sigma_r)$ be the ratio of the observed width $w_{\mathrm{M}}(\sigma_r)$ to σ_r. The width ratio $w'_{\mathrm{M}}(\sigma_r)$ is constant regardless of σ_r, that is,

$$w'_{\mathrm{M}}(\sigma_r) = \frac{w_{\mathrm{M}}(\sigma_r)}{\sigma_r} = w'_{\mathrm{M}_c}, \tag{10.24}$$

where $w'_{\mathrm{M}_c} \approx 1.0$ when $\gamma_{23} = 1$ in the formulation of Eq. (10.2). Similarly, we define the height measure $h_{\mathrm{R}}(\sigma_r;\sigma_f)$ of the single-scale filter response as the peak response $h_{\mathrm{R}}(\sigma_r;\sigma_f) = \mathcal{S}_{line}\{\ell_{line}(0,0,z;\sigma_r);\sigma_f\}$ and the width measure $w_{\mathrm{R}}(\sigma_r,\sigma_f)$ as the distance $\sqrt{x_0^2 + y_0^2}$ from the z-axis to the circular locus where $\mathcal{S}_{line}\{\ell_{line}(x_0, y_0, z;\sigma_r);\sigma_f\}$ gives the half of the maximum response, $\frac{h_{\mathrm{M}_c}}{2}$. To compare the widths of the filter response and the original profile, we also introduce the width measure $w_{\mathrm{L}}(\sigma_r)$ of the original line image as the distance $\sqrt{x_0^2 + y_0^2}$ from the z-axis to the circular locus where $\ell_{line}(x_0, y_0, z;\sigma_r)$ gives half of $\ell_{line}(0,0,z;\sigma_r)$. While σ_r is introduced for the convenience of generating line profiles, $w_{\mathrm{L}}(\sigma_r)$ is for the convenience of comparing the widths of various profile shapes.

Figure 10.10(a) and 10.10(b) show the variations in the height and width measures. Figure 10.10(a) gives the plots of $h_{\mathrm{R}}(\sigma_r;\sigma_f)$ at three values of σ_f and $h_{\mathrm{M}}(\sigma_r)$, and Fig. 10.10(b) shows the plots of $w_{\mathrm{R}}(\sigma_r;\sigma_f)$ at three values of σ_f, $w_{\mathrm{M}}(\sigma_r)$, and $w_{\mathrm{L}}(\sigma_r)$. The width measure of the multiscale response is proportional to that of the original line image. In the case of the line image $\ell_{line}(\vec{x};\sigma_r)$ with a Gaussian cross section, $w_{\mathrm{L}}(\sigma_r) \approx 0.9 w_{\mathrm{M}}(\sigma_r)$. Although the filter responses make the lines a little thinner than the original lines, the multiscale line-filter can be designed so that the width of its responses becomes approximately proportional to the original one.

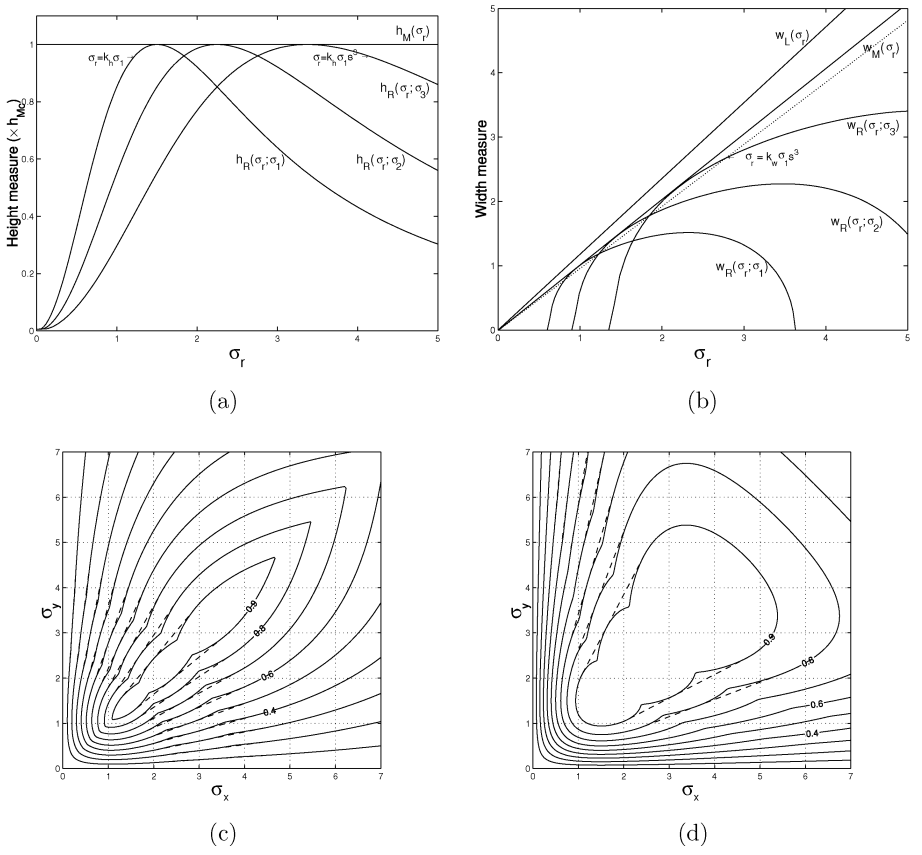


Figure 10.10: Height and width measures of filter responses in multiscale integration with $\sigma_i = s^{i-1}\sigma_1$ ($\sigma_1 = 1.5, s = 1.5, i = 1, 2, 3$). The height measure is normalized so that h_{M_c} is one. (a) Height measures $h_{\mathrm{M}}(\sigma_r) = h_{\mathrm{M}_c}$ and $h_{\mathrm{R}}(\sigma_r; \sigma_i)$ for $\ell_{line}(\vec{x}; \sigma_r)$. (b) Width measures $w_{\mathrm{M}}(\sigma_r) = w_{\mathrm{M}_c}$, $w_{\mathrm{L}}(\sigma_r)$, and $w_{\mathrm{R}}(\sigma_r; \sigma_i)$ for $\ell_{line}(\vec{x}; \sigma_r)$ with $\gamma_{23} = 1$. (c) Height measures for $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$ with $\gamma_{23} = 1$. Solid lines denote the height measures for the discrete scales. Dashed lines denote the height measures for the continuous scales from σ_1 to σ_3. (d) Height measures for $\ell_{elliptic}(\vec{x}; \sigma_x, \sigma_y)$ with $\gamma_{23} = 0.5$.

10.3.3 Multiscale Responses of Discrete Scale Integration

We assumed continuous scales for multiscale integration in the previous subsection. The response at each scale, however, has to be computed at discrete

values of σ_f. In 3-D image filtering, a large amount of computation is necessary to obtain the filter response at each value of σ_f. The maximum and minimum values of σ_f can be essentially determined on the basis of the width range of the anatomical structure of interest. The interval of σ_f should be sufficiently small for the filter response to work uniformly for every line width within the width range, while it is desirable that it should be large enough to minimize the amount of computation required. Thus, we need to determine the minimum number of discrete samples of σ_f which satisfy the following conditions:

1. The height measure of the response should be approximately constant within the width range.
2. The width measure of the response should be approximately proportional to the original one within the width range.

Given the discrete samples of σ_f and the assumption of the cross-section shape (here, we use the Gaussian cross section), the accuracy of the approximation can be estimated. Let $\sigma_i = s^{i-1}\sigma_1$ $(i = 1, 2, \ldots, n)$ be discrete samples of σ_f, where σ_1 is the minimum scale and s is the scale factor determining the sampling interval of σ_f. The multiscale filter response using the discrete samples of σ_f is given by

$$\mathcal{M}_{line}\{f\} = \max_{1 \le i \le n} \mathcal{S}_{line}\{f; \sigma_i\}. \tag{10.25}$$

Similarly, the multiscale filter response using the discrete samples of σ_f for the line image $\ell_{line}(\vec{x}, \sigma_r)$ is given by

$$\mathcal{M}_{line}\{\ell_{line}(\vec{x}, \sigma_r)\} = \max_{1 \le i \le n} \mathcal{S}_{line}\{\ell_{line}(\vec{x}, \sigma_r); \sigma_i\}. \tag{10.26}$$

Given the scale factor s, we can determine $h_{\mathrm{M}_{min}}$ and k_p satisfying

$$h_{\mathrm{M}_{min}} = h_{\mathrm{R}}(k_p\sigma_i; \sigma_i) = h_{\mathrm{R}}(k_p s\sigma_i; \sigma_i), \tag{10.27}$$

where $\sigma_i = s^{i-1}\sigma_1$ $(i = 1, 2, \ldots, n)$, $h_{\mathrm{M}_{min}}$ is the minimum of the height measure of the multiscale response $\mathcal{M}_{line}\{\ell_{line}(\vec{x}, \sigma_r)\}$ within the range $k_p\sigma_1 \le \sigma_r \le k_p s^n\sigma_1$, and the minimum is taken at $\sigma_r = k_p s^i\sigma_1$ $(i = 0, 1, 2, \ldots n)$ (Fig. 10.10(a)). $h_{\mathrm{M}_{min}}$ can be regarded as a function of the scale factor s. The height measure of the multiscale response should be sufficiently close to h_{M_c} within the width range of interest. The values of $h_{\mathrm{M}_{min}}$ and k_p at typical scale factors are summarized in

Table 10.2: Height measure $h_{\mathrm{M}}(\sigma_r)$ and width ratio $w'_{\mathrm{M}}(\sigma_{\mathrm{r}})$ minima in multi-scale integration at discrete scales using typical scale factors, and σ_{r} where the minima are taken

Scale factor s	Min. height $h_{\mathrm{M}_{min}}$	at $\sigma_{\mathrm{r}} = k_p\sigma_i$	Min. width ratio $w'_{\mathrm{M}_{min}}$	at $\sigma_{\mathrm{r}} = k_w\sigma_i$
$s \to 1$	$h_{\mathrm{M}_{min}} \to h_{\mathrm{M}_c}$	$k_p \to 1$	$w'_{\mathrm{M}_{min}} \to w'_{\mathrm{M}_c}$	$k_w \to 0.65$
$s = 1.2$	$h_{\mathrm{M}_{min}} \approx 0.99h_{\mathrm{M}_c}$	$k_p \approx 0.92$	$w'_{\mathrm{M}_{min}} \approx 0.99w'_{\mathrm{M}_c}$	$k_w \approx 0.59$
$s = \sqrt{2}$	$h_{\mathrm{M}_{min}} \approx 0.97h_{\mathrm{M}_c}$	$k_p \approx 0.84$	$w'_{\mathrm{M}_{min}} \approx 0.97w'_{\mathrm{M}_c}$	$k_w \approx 0.56$
$s = 1.5$	$h_{\mathrm{M}_{min}} \approx 0.96h_{\mathrm{M}_c}$	$k_p \approx 0.82$	$w'_{\mathrm{M}_{min}} \approx 0.96w'_{\mathrm{M}_c}$	$k_w \approx 0.55$
$s = 2.0$	$h_{\mathrm{M}_{min}} \approx 0.89h_{\mathrm{M}_c}$	$k_p \approx 0.71$	$w'_{\mathrm{M}_{min}} \approx 0.88w'_{\mathrm{M}_c}$	$k_w \approx 0.50$

Table 10.2. When $s = 1.5$, $h_{\mathrm{M}_{min}} \approx 0.96h_{\mathrm{M}_c}$, which means that the deviation from the continuous case is less than 4%.

With regard to the width measure of the filter response, given the discrete samples of σ_f and the assumption of the profile shape, the accuracy of this approximation can also be estimated. Given the scale factor s, we can determine $w'_{\mathrm{M}_{min}}$ and k_w satisfying

$$w'_{\mathrm{M}_{min}} = \frac{w_{\mathrm{R}}(k_w\sigma_i;\sigma_i)}{k_w\sigma_i} = \frac{w_{\mathrm{R}}(k_w s\sigma_i;\sigma_i)}{k_w s\sigma_i}, \tag{10.28}$$

where $\sigma_i = s^{i-1}\sigma_1$ $(i = 1, 2, \ldots, n)$, $w'_{\mathrm{M}_{min}}$ is the minimum of the ratio of $w_{\mathrm{M}}(\sigma_r)$ to σ_r within the range $k_w\sigma_1 \le \sigma_r \le k_w s^n\sigma_1$, and the minimum is taken at $\sigma_r = k_w s^i\sigma_1$ $(i = 0, 1, 2, \ldots, n)$ (Fig. 10.10(b)). $w'_{\mathrm{M}_{min}}$ can be regarded as a function of the scale factor s. $w'_{\mathrm{M}}(\sigma_r)$ should be sufficiently close to w'_{M_c} within the width range of interest. The values of $w'_{\mathrm{M}_{min}}$ and k_w at typical scale factors are summarized in Table 10.2. When $s = 1.5$, $w'_{\mathrm{M}_{min}} \approx 0.96w'_{\mathrm{M}_c}$. When the parameters for the discrete scales of σ_f are $s = 1.5$ and $n = 3$, the ranges of deviation within 4% for the height and the width measures are $0.55\sigma_1 \le \sigma_r \le 1.86\sigma_1$ and $0.82\sigma_1 \le \sigma_r \le 2.77\sigma_1$, respectively. The range of deviation within 4% for the width measure is shifted to a smaller σ_r than that for the height measure. As a result, the range of deviation of less than 4% for both the height and the width measures is $0.82\sigma_1 \le \sigma_r \le 1.86\sigma_1$.

We now extend the experimental analysis of the multiscale integration to the response to $\ell_{elliptic}(\vec{x};\sigma_x,\sigma_y)$ shown in Eq. (10.21). We define the hight measure of $\ell_{elliptic}(\vec{x};\sigma_x,\sigma_y)$ as $h_{\mathrm{R}_{elliptic}}(\sigma_x,\sigma_y;\sigma_f) = \mathcal{S}_{line}\{\ell_{elliptic}(0,0,z;\sigma_x,\sigma_y);\sigma_f\}$. Fig. 10.10(c) and 10.10(d) show the multiscale integration of the responses at continuous and discrete scales with $\sigma_1 = 1.5$, $s = 1.5$ and $n = 3$ for $\gamma_{23} = 1$ and

$\gamma_{23} = 0.5$, respectively. The multiscale integration of these discrete scales gives a good approximation of that of the continuous scales.

10.4 Description and Quantification

In the previous sections, the enhancement of the local structures based on the eigenvalues of the Hessian matrix was discussed. In this section, we further combine the gradient vector with the Hessian matrix to perform explicit detection, localization, and description of the local structures. Especially, we focus on the line and sheet structures. The methods are formulated as a 3-D extension of 2-D line description presented in [36]. The 3-D line model consists of the *medial axes* of lines and the *cross-sectional shape* associated with each point on these axes, while the 3-D sheet model consists of the *medial surfaces* of sheets and the *width* associated with each point on these surfaces. The medial axes and medial surfaces are detected and localized by fully utilizing formal analyses of 3-D second-order local intensity structures based on the gradient vector and the Hessian matrix.

The following is an overview of the method:

Step 1: Existing filtering techniques for line and sheet enhancement are used to extract the initial regions, which should include all potential medial axes and surfaces [7, 11]. These are then used as initial values for the subsequent subvoxel edge localization. The candidate regions, which should include all potential line and sheet regions, are also extracted.

Step 2: The medial axes and surfaces are extracted using local second-order approximation given by the gradient vector and Hessian matrix. The eigenvectors of the Hessian matrix define the moving frames on medial axes/surfaces. After this, the moving frames are embedded in a 3-D image such that each point within the candidate regions is directly related to its corresponding moving frame.

Step 3: Subvoxel edge localization of the region boundaries is carried out using adaptive 3-D directional derivatives, whose directions are adaptively changed depending on the moving frame, to accomplish accurate segmentation, model recovery, and quantification.

In the following, we begin with a description of Step 2 of the method since Step 1 has been already described in the previous sections.

10.4.1 Medial Axis and Surface Detection

Let $f(\vec{x})$ be an intensity function of a volume, where $\vec{x} = (x, y, z)$, and $f(\vec{x};\sigma)$ be its Gaussian smoothed volume with standard deviation σ. The second-order approximation of $f(\vec{x};\sigma)$ around $\vec{x}_0$ is given by

$$f_{\mathrm{II}}(\vec{x};\sigma) = f_0 + (\vec{x}-\vec{x}_0)^\top \nabla f_0 + \frac{1}{2}(\vec{x}-\vec{x}_0)^\top \nabla^2 f_0 (\vec{x}-\vec{x}_0), \qquad (10.29)$$

where $f_0 = f(\vec{x}_0)$, $\nabla f_0 = \nabla f(\vec{x}_0)$, and $\nabla^2 f_0 = \nabla^2 f(\vec{x}_0)$. Thus, the second-order structures of local intensity variations around each point of a volume can be described by the original intensity, the gradient vector, and the Hessian matrix.

The gradient vector of Gaussian smoothed volume $f(\vec{x};\sigma)$ is defined as

$$\nabla f(\vec{x};\sigma) = (f_x(\vec{x};\sigma), f_y(\vec{x};\sigma), f_z(\vec{x};\sigma))^\top, \qquad (10.30)$$

where partial derivatives of $f(\vec{x};\sigma)$ are represented as $f_x(\vec{x};\sigma) = \frac{\partial}{\partial x} f(\vec{x};\sigma)$, $f_y(\vec{x};\sigma) = \frac{\partial}{\partial y} f(\vec{x};\sigma)$, and $f_z(\vec{x};\sigma) = \frac{\partial}{\partial z} f(\vec{x};\sigma)$.

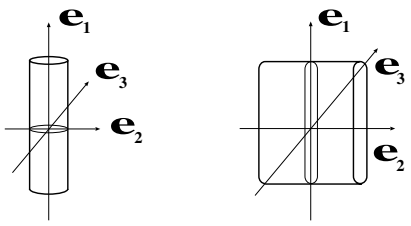
The Hessian matrix of Gaussian smoothed volume $f(\vec{x};\sigma)$ is given by

$$\nabla^2 f(\vec{x};\sigma) = \begin{bmatrix} f_{xx}(\vec{x};\sigma) & f_{xy}(\vec{x};\sigma) & f_{xz}(\vec{x};\sigma) \\ f_{yx}(\vec{x};\sigma) & f_{yy}(\vec{x};\sigma) & f_{yz}(\vec{x};\sigma) \\ f_{zx}(\vec{x};\sigma) & f_{zy}(\vec{x};\sigma) & f_{zz}(\vec{x};\sigma) \end{bmatrix}, \qquad (10.31)$$

where partial second derivatives of $f(\vec{x};\sigma)$ are represented as $f_{xx}(\vec{x};\sigma) = \frac{\partial^2}{\partial x^2} f(\vec{x};\sigma)$, $f_{yz}(\vec{x};\sigma) = \frac{\partial^2}{\partial y \partial z} f(\vec{x};\sigma)$, and so on.

Let the eigenvalues of $\nabla^2 f(\vec{x};\sigma)$ be $\lambda_1, \lambda_2, \lambda_3$ ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) and their corresponding eigenvectors be $\vec{e}_1, \vec{e}_2, \vec{e}_3$ ($|\vec{e}_1| = |\vec{e}_2| = |\vec{e}_3| = 1$), respectively. For the ideal line, $\vec{e}_1$ is expected to give its tangential direction and both $|\lambda_2|$ and $|\lambda_3|$, directional second derivatives orthogonal to $\vec{e}_1$, should be large on its medial axis, while $\vec{e}_3$ is expected to give the orthogonal direction of a sheet and only $|\lambda_3|$ should be large on its medial surface (Fig. 10.11). Here, structures of interest are assumed to be brighter than surrounding regions.

The initial regions obtained in Step 1 are searched for medial axes and surfaces, which are detected based on the second-order approximation of $f(\vec{x};\sigma)$ The medial axis and surface extraction is based on a formal analysis of the second-order 3D local intensity structure. Here, σ_f is the filter scale used in



(a) (b)

Figure 10.11: Line and sheet models with the eigenvectors of the Hessian matrix. (a) Line. (b) Sheet.

medial axis/surface detection, and we assume that the width range of structures of interest is around the width at which the filter with σ_f gives the peak response (see [7] and [11] for detailed discussions).

10.4.1.1 Line Case—Medial Axis Detection

We assume that the tangential direction is given by $\vec{e}_1$ at the voxel around the medial axis. The 2-D intensity function, $c(\vec{u})$ ($\vec{u} = (u, v)^\top$), on the cross-sectional plane of $f(\vec{x}; \sigma_f)$ orthogonal to $\vec{e}_1$, should have its peak on the medial axis. The second-order approximation of $c(\vec{u})$ is given by

$$c(\vec{u}) = f(\vec{x}_0; \sigma_f) + \vec{u}^\top \nabla c_0 + \frac{1}{2}\vec{u}^\top \nabla^2 c_0 \vec{u}, \tag{10.32}$$

where $u\vec{e}_2 + v\vec{e}_3 = \vec{x} - \vec{x}_0$, $\nabla c_0 = (\nabla f \cdot \vec{e}_2, \nabla f \cdot \vec{e}_3)^\top$ (∇f is the gradient vector, that is, $\nabla f(\vec{x}_0; \sigma_f)$), and

$$\nabla^2 c_0 = \begin{bmatrix} \lambda_2 & 0 \\ 0 & \lambda_3 \end{bmatrix}. \tag{10.33}$$

$c(\vec{u})$ should have its peak on the medial axis of the line. The peak is located at the position satisfying

$$\frac{\partial}{\partial u} c(\vec{u}) = 0 \quad \text{and} \quad \frac{\partial}{\partial v} c(\vec{u}) = 0. \tag{10.34}$$

By solving Eq. (10.34), we have the offset vector, $\vec{p} = (p_x, p_y, p_z)^\top$, of the peak position from $\vec{x}_0$ given by

$$\vec{p} = s\vec{e}_2 + t\vec{e}_3, \tag{10.35}$$

where $s = -\frac{\nabla f \cdot \vec{e}_2}{\lambda_2}$ and $t = -\frac{\nabla f \cdot \vec{e}_3}{\lambda_3}$. For the medial axis to exist at the voxel $\vec{x}_0$, the peak of $c(\vec{u})$ needs to be located in the territory of voxel $\vec{x}_0$. Thus, the medial axis is detected only if $|p_x| \leq \frac{1}{2}$, $|p_y| \leq \frac{1}{2}$, and $|p_z| \leq \frac{1}{2}$. By combining the voxel position $\vec{x}_0$ and offset vector $\vec{p}$, the medial axis is localized at subvoxel resolution.

10.4.1.2 Sheet Case—Medial Surface Detection

We assume that the direction of the surface normal is given by $\vec{e}_1$ at the voxel around the medial surface. The 1-D intensity function, $c(v)$, which is the profile of $f(\vec{x}; \sigma_f)$ along $\vec{e}_3$, should have its peak on the medial surface. The second-order approximation of $c(v)$ is given by

$$c(v) = f(\vec{x}_0; \sigma_f) + vc_0' + \frac{1}{2}v^2 c_0'', \tag{10.36}$$

where $v\vec{e}_3 = \vec{x} - \vec{x}_0$, $c_0' = \nabla f \cdot \vec{e}_3$, and $c_0'' = \lambda_3$. $c(v)$ should have its peak on the medial surface of the sheet. The peak is located at the position satisfying

$$\frac{\mathrm{d}}{\mathrm{d}v}c(v) = 0. \tag{10.37}$$

By solving Eq. (10.37), we have the offset vector, $\vec{p}$, of the peak position from $\vec{x}_0$ given by

$$\vec{p} = t\vec{e}_3, \tag{10.38}$$

where $t = -\frac{\nabla f \cdot \vec{e}_3}{\lambda_3}$. The medial surface is detected only if $|p_x| \leq \frac{1}{2}$, $|p_y| \leq \frac{1}{2}$, and $|p_z| \leq \frac{1}{2}$.

10.4.1.3 Embedding Moving Frames

The moving frame is defined by the voxel position $\vec{x}_0$, the offset vectors $\vec{p}$, and the eigenvectors $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ at each detected point of a medial axis or surface. In order to perform the subsequent processes based on moving frames, each voxel within the candidate regions obtained in Step 1 needs to be related to the moving frame. First, we find the correspondences between each voxel and one of the detected points of a medial axis or surface. Once these correspondences are found, each voxel is directly related to its corresponding moving frame. To find the correspondences, we use the Voronoi tessellation of the detected points.

The territory of the detected point in the Voronoi tessellation can be regarded as the set of voxels to which each discrete moving frame is applied. This process is identical in both the line and sheet cases.

10.4.2 Subvoxel Edge Localization and Width Measurement

An adaptive directional second derivative is applied at each voxel based on its corresponding moving frame. The directional derivative is taken along the perpendicular from the voxel to the medial axis or surface. The zero-crossing points of the directional second derivatives are localized at subvoxel resolution to determine the precise region boundaries and quantitate the widths.

At every voxel within the candidate regions, the directional second derivative is calculated depending on its corresponding moving frame. This spatially variable directional derivative is written as

$$f''_{line}(\vec{x};\sigma_{\mathrm{e}}) = \vec{r}(\vec{x})^{\top}\nabla^2 f(\vec{x};\sigma_{\mathrm{e}})\vec{r}(\vec{x}), \tag{10.39}$$

where $\vec{r}(\vec{x})$ is the unit vector whose direction is parallel to the perpendicular from the voxel position $\vec{x}$ to the straight line defined by the origin and the medial axis direction of the moving frame. The foot of the perpendicular can be regarded as the corresponding axis position. The origin is given by the voxel position of the medial axis point and the offset vector $\vec{p}$. σ_{e} is the filter scale used in the edge localization; it is desirable that σ_{e} be small compared to the line width for accurate edge localization.

After the adaptive derivatives have been calculated at all the voxels, subvoxel edge localization is carried out at every voxel in the candidate regions. Let $\vec{o}_{\mathrm{a}}$ be the foot of the perpendicular on the axis. Let $\vec{r}_{\mathrm{a}}$ be the direction from $\vec{o}_{\mathrm{a}}$ to the voxel position $\vec{x}_{\mathrm{a}}$. For each voxel, we reconstruct the profiles originating from $\vec{o}_{\mathrm{a}}$ in the directions $\vec{r}_{\mathrm{a}}$ and $-\vec{r}_{\mathrm{a}}$ for $f''_{line}(\vec{x};\sigma_{\mathrm{e}})$ and the initial regions (which we specify as $b_{line}(\vec{x})$) obtained in Step 1. The edges are then localized in both directions and the width is calculated as the distance between the two edge locations. The profile is reconstructed at subvoxel resolution by using a trilinear interpolation for $f''_{line}(\vec{x};\sigma_{\mathrm{e}})$ and a nearest-neighbor interpolation for $b_{line}(\vec{x})$.

Let $f''(s)$ be the profiles of $f''_{line}(\vec{x};\sigma_{\mathrm{e}})$ along the directions $\vec{r}_{\mathrm{a}}$ from $\vec{o}_{\mathrm{a}}$. Let $b(s)$ be the profiles of $b_{line}(\vec{x})$. Here, r denotes the position from the foot of

the perpendicular on the axis. The localization of edges consists of two steps; finding the initial point for the subsequent search using $b(s)$, and then searching for the zero-crossing of $f''(s)$. The initial point, p_0, is given by s of the first encountered point satisfying $b(s) = 0$, starting the search from $s = 0$, that is, the axis point, to the direction $\vec{r}_{\mathrm{a}}$. Given the initial point of the search, if $f''(p_0) < 0$, search outbound from the axis point along the profile for the zero-crossing position p; otherwise, search inbound. After the zero-crossing position q in the opposite direction $-\vec{r}_{\mathrm{a}}$ is similarly determined, the width (diameter) is given by $|p - q|$.

10.4.2.1 Sheet Case

At every voxel within the candidate regions, the directional second derivative is taken orthogonal to both $\vec{e}_1$ and $\vec{e}_2$, that is, along $\vec{e}_3$, in its corresponding moving frame. This spatially variable directional derivative is written as

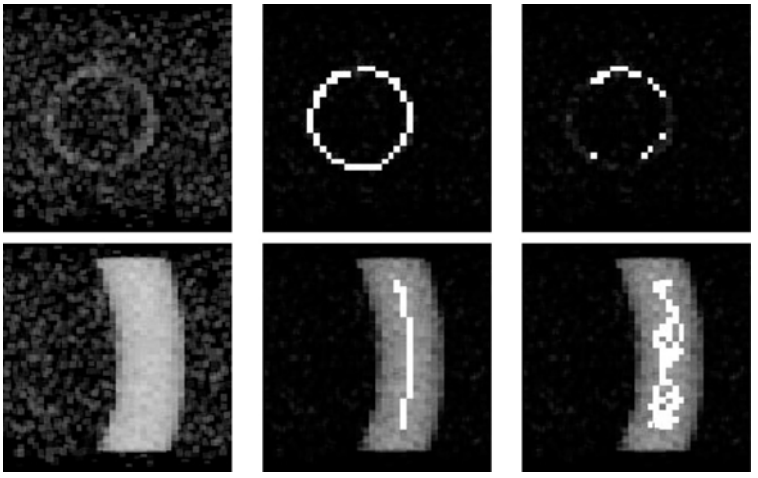
$$f''_{sheet}(\vec{x}; \sigma_{\mathrm{e}}) = \vec{r}(\vec{x})^{\top} \nabla^2 f(\vec{x}; \sigma_{\mathrm{e}}) \vec{r}(\vec{x}), \tag{10.40}$$

where $\vec{r}(\vec{x})$ is the unit vector whose direction is parallel to the medial surface normal of the moving frame. Using a method analogous to that employed in the line case, the profiles $f''(s)$ and $b(s)$ of $f''_{sheet}(\vec{x}; \sigma_e)$ and $b_{sheet}(\vec{x})$ are reconstructed for the directions $\vec{r}$, respectively. These profiles are then used to determine the edge locations p and q in the two directions $\vec{r}$ and $-\vec{r}$, respectively, and finally the width (thickness) is obtained as $|p - q|$.

10.4.3 Simulational Evaluation of Medial Axis Detection

We evaluated the medial axis detection performance using synthesized 3-D images of lines with pill-box cross-sections. A simulated partial volume effect was incorporated when synthesizing the images. We focused on the effects of the filter scale σ_f used in medial axis detection on the detection various widths of line structures.

Synthesized 3-D images of a line with a circular axis were generated. The diameter of the line, D, was varied between 2.0 and $8\sqrt{2}(\simeq 11.3)$ voxels. The radius of the circular axis was proportional to D (we used $4 \times D$). Gaussian noise with 25% standard deviation in the intensity height of pill-box cross-sections was added to the images. After line and sheet enhancement filtering with



(a) (b) (c)

Figure 10.12: Medial axis detection from synthesized 3-D images. All units are voxels. (a) Volume-rendered images of original synthesized 3D images with Gaussian noise. Upper: $D = 2.0$. Lower: $D = 11.3$. (b) Examples of successful axis detection where the filter scale σ_f is appropriate for the line diameter D. The detected axis points are shown as bright points. Upper: $D = 2.0$, $\sigma_f = 1.4$. Lower: $D = 11.3$, $\sigma_f = 4.0$. (c) Examples of undesirable axis detection. Upper: $D = 2.0, \sigma_f = 4.0$. When diameter D is smaller than that appropriate for the filter scale σ_f, many true axis points are overlooked. Lower: $D = 11.3, \sigma_f = 1.4$. When D is larger than that appropriate for σ_f, many false axis points are detected.

integration of scales appropriate for the line diameter [7, 11], the candidate regions were extracted by thresholding and extracting large connective components. The medial axis points were detected within these regions using the procedures described in Section 10.4.1 with two values for the filer scale σ_f, $\sqrt{2}$ ($\simeq 1.4$) and 4.0 voxels. The same candidate regions were used for both values of σ_f.

Figure 10.12 shows the volume rendering of the synthesized 3-D images and typical axis detection results. The detection was successful using appropriate combinations of line diameter D and filter scale σ_f (Fig. 10.12(b)). Many axis points are overlooked when the filter scale is larger than appropriate, while a number of false detections are made when the filter scale is smaller than appropriate (Fig. 10.12(c)).

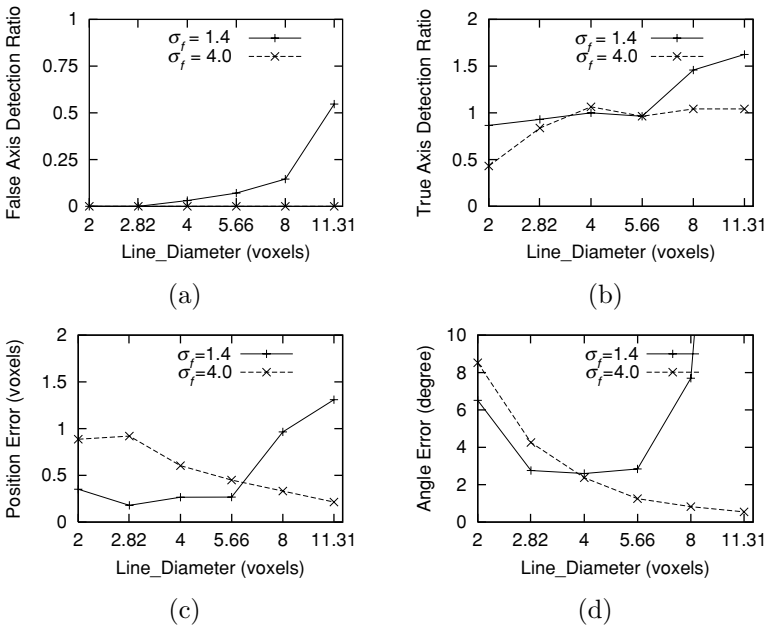


Figure 10.13: Performance evaluation of medial axis detection. See text for the definitions of true and false detections. (a) False positive detection ratio, which is the ratio of the number of false detections to all the detections. The ratio was zero for all the line diameters at $\sigma_f = 4.0$ voxels. (b) True positive detection ratio, which is the ratio of the number of true detections to all the analytically determined points. (c) Average position error of axis points regarded as true detections. The distance between detected points and analytically determined points was used as the error. (d) Average angle error of the directions of axis points regarded as true detections.

Figure 10.13 shows the performance evaluation results. Detected axis points were evaluated by comparing them with analytically determined axis points. We regarded a detected point as a true detection if the distance between its position and one of the analytically determined points was within two voxels; otherwise, detected points were regarded as false. The false and true positive detection ratios are shown in Figs. 10.13(a) and 10.13(b); the plots verify the observations in Fig. 10.12. The positions and directions of the detected axis points regarded as true detections were compared with analytically determined ones (Figs. 10.13(c) and 10.13(d)). These graphs clarify the effect of σ_f on accurate and reliable axis detection.

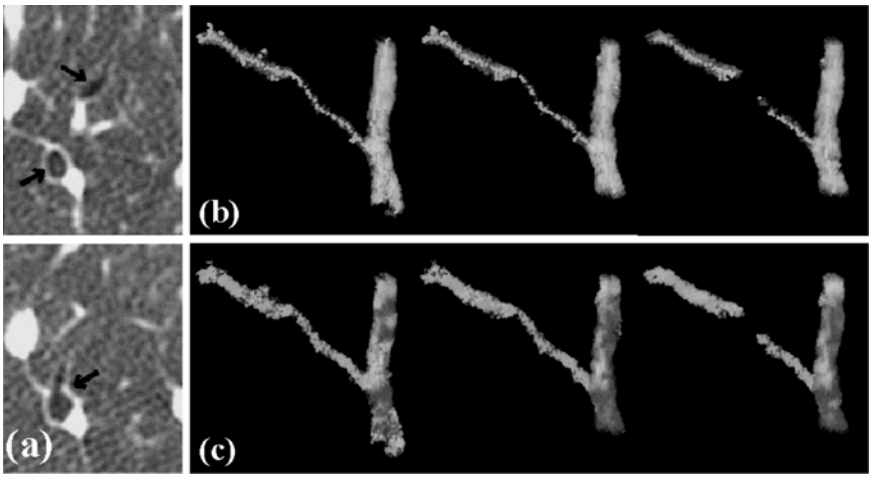


Figure 10.14: Diameter estimation of bronchial airways from CT images. (a) Original CT images. The bronchial airway regions, which are darker than the surrounding structures, are shown by arrows. (b) Detection of medial axes at three different scales. Left: $\sigma_f = 1.4$ voxels. Middle: $\sigma_f = 2.0$ voxels. Right: $\sigma_f = 2.8$ voxels. (c) Diameter estimation at the three different scales. (A color version of this figure will appear on the CD that accompanies the volume.)

10.4.4 Examples

10.4.4.1 Bronchi Diameter Quantification from 3-D CT Data

The line width quantification method was applied to chest CT images taken by a helical CT scanner to determine the diameters of bronchi. The original voxel dimensions were $0.29 \times 0.29 \times 1.0$ (mm^3). In order to make the voxel isotropic, sinc interpolation was applied along the z-direction. The volume size used in the experiment was $90 \times 70 \times 80$ (voxels) after interpolation.

Figure 10.14(a) shows the original CT images. After the initial region extraction by thresholding the line filtered images, the medial axis was detected using $\sigma_f = \sqrt{2}, 2$, and $2\sqrt{2}$ voxels. Figure 10.14(b) shows the results of axis detection at the three different scales. Note that the axis points of thin structures were detected only at the smaller two scales while those of large structures (the right segment) were stably extracted at the larger two scales. Figure 10.14(c) shows the results of diameter estimation using $\sigma_e = 1.2$ voxels based on the medial axes at these three scales.

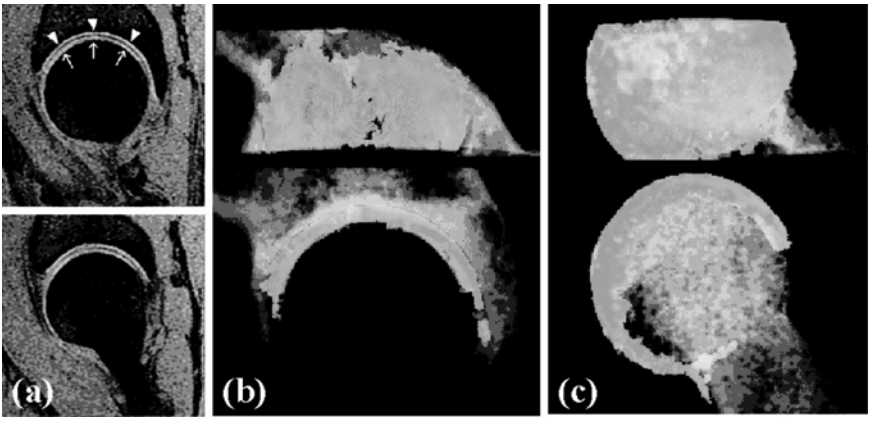


Figure 10.15: Thickness estimation of cartilages from MR images. (a) Original MR images. The acetabular (pelvic side) cartilages are shown by arrowheads and the femoral head cartilages by arrows. (b) Thickness distribution of acetabular cartilages. The bone regions are volume-rendered in white. (c) Thickness distribution of femoral cartilages. (A color version of this figure will appear on the CD that accompanies the volume.)

10.4.4.2 Hip Joint Cartilage Thickness Quantification from 3-D MR Data

The sheet width quantification method was applied to MR images of a hip joint [37, 38] to determine the thickness of hip joint cartilage. The original voxel dimensions were $0.62 \times 0.62 \times 1.5$ (mm^3). Sinc interpolation was applied along the z-direction to make the voxel isotropic, and then further applied along all the three directions to make the resolution double. The resultant sampling pitch was 0.31 (mm) in all the three directions. The volume size used in the experiment was $256 \times 256 \times 100$ (voxels) after interpolation.

As shown in Fig. 10.15(a), cartilages are thin structures; thickness distributions are considered to be particularly important in the diagnosis of joint diseases. The initial cartilage regions were extracted from the enhanced images by the sheet filter. The medial surfaces were extracted using $\sigma_f = 1.4$ voxels. Figures 10.15(b) and 10.15(c) show the results of thickness distribution estimated using $\sigma_e = 1.2$ voxels. We also obtained the thickness distributions using $\sigma_e = 1.0$ voxel for comparison purposes. The average thickness estimated using $\sigma_e = 1.2$ voxels was $T_f = 4.24$ voxels and $T_a = 3.50$ voxels for the femoral and

acetabular cartilages, respectively, compared with $T_{\mathrm{f}} = 4.16$ voxels and $T_{\mathrm{a}} = 3.39$ voxels using $\sigma_{\mathrm{e}} = 1.0$ voxel.

In the related work [10], hip joint cartilages were assumed to be distributed on a sphere approximating the femoral head. The user needs to specify the center of the sphere, and the cartilage thickness is then estimated along radial directions from the specified center. The method applied here does not use the sphere assumption, and thus can potentially be applied to badly deformed hip joints as well as to articular cartilages of other joints.

10.5 Analysis for Sheet Width Quantification Accuracy

In this section, we present a systematic approach to the accuracy validation of width quantification. Especially, we investigate inherent limits on the accuracy of sheet width measurement described in the previous section arising from finite resolution. We focus on MR imaged structure, and especially address the question as to how the accuracy depends on the orientation of sheet structures when a voxel shape is anisotropic in MR imaging. In the following, a theoretical procedure for ascertaining the inherent limits on the accuracy of sheet width (thickness) measurement in MR images is presented.

10.5.1 Mathematical Modeling of MR Imaging and Width Measurement Processes

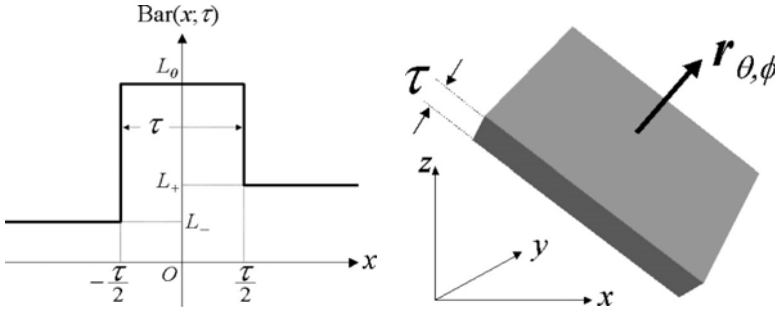
10.5.1.1 Modeling a Sheet Structure

A 3-D sheet structure orthogonal to the x-axis is modeled as

$$s_0(\vec{x}; \tau) = \mathrm{Bar}(x; \tau), \tag{10.41}$$

where $\vec{x} = (x, y, z)^\top$, and

$$\mathrm{Bar}(x; \tau) = \begin{cases} L_- & x < -\frac{1}{2}\tau \\ L_0 & -\frac{1}{2}\tau \le x \le \frac{1}{2}\tau, \\ L_+ & x > \frac{1}{2}\tau \end{cases} \tag{10.42}$$



(a) (b)

Figure 10.16: Modeling 3-D sheet structures. (a) Bar profile of MR values along sheet normal direction with thickness τ. L_0, L_-, and L_+ denote sheet object, left-side, and right-side background levels, respectively. (b) 3-D sheet structures with thickness τ and normal orientation $\vec{r}_{\theta,\phi}$. (© 2004 IEEE)

in which τ represents the thickness (width) of the sheet. L_0, L_-, and L_+ are the MR signal intensities of the sheet and both sides of backgrounds, respectively (Fig. 10.16(a)). Let (θ, ϕ) be a pair of latitude and longitude which represents the normal orientation of the sheet given by

$$\vec{r}_{\theta,\phi} = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)^\top. \tag{10.43}$$

The 3-D sheet structure with orientation $\vec{r}_{\theta,\phi}$ is written as

$$s(\vec{x}; \tau, \vec{r}_{\theta,\phi}) = s_0(\vec{x}'; \tau), \tag{10.44}$$

where $\vec{x}' = R_{\theta,\phi}\vec{x}$, in which $R_{\theta,\phi}$ denotes a 3×3 matrix representing rotation which enables the normal orientation of the sheet $s_0(\vec{x}; \tau)$, i.e. the x-axis, correspond to $\vec{r}_{\theta,\phi}$ (Fig. 10.16(b)).

10.5.1.2 Modeling MR Image Acquisition

The 1-D point spread function (PSF) of MR images [39] is given by

$$m(x; \Delta_x) = \frac{1}{N_x} \frac{\sin(\pi \frac{x}{\Delta_x})}{\sin(\pi \frac{x}{N_x \Delta_x})}, \tag{10.45}$$

where N_x is the number of samples in the frequency domain, and Δ_x represents the sampling interval in the spatial domain. Eq. (10.45) is well approximated

[40] by

$$m(x;\Delta_x) = \mathrm{Sinc}\left(x;\frac{1}{\Delta_x}\right). \tag{10.46}$$

where

$$\mathrm{Sinc}(x;w) = \frac{\sin(\pi wx)}{\pi wx}. \tag{10.47}$$

The 3-D PSF is given by

$$m(\vec{x};\Delta_x,\Delta_y,\Delta_z) = m(x;\Delta_x)\cdot m(y;\Delta_y)\cdot m(z;\Delta_z), \tag{10.48}$$

where Δ_x, Δ_y, and Δ_z are sampling intervals along the x-axis, y-axis, and z-axis, respectively.

In actual MR imaging, the magnitude operator is applied to the complex number obtained at each voxel by FFT reconstruction, whose effects are not negligible [41]. Thus, the MR image of the sheet structure with orientation $\vec{r}_{\theta,\phi}$ and thickness τ is given by

$$f(\vec{x}) = |s(\vec{x};\tau,\vec{r}_{\theta,\phi}) * m(\vec{x};\Delta_x,\Delta_y,\Delta_z)|, \tag{10.49}$$

where $*$ denotes the convolution operation.

10.5.1.3 Thickness Determination Procedure

In this chapter, we restrict the scope of our investigation to the sheet model described in section 10.5.1 (Fig. 10.16), that is, a sheet structure with constant thickness τ and orientation $\vec{r}_{\theta,\phi}$. We define the thickness measured from the MR imaged sheet structure as the distance between both sides of image edges along the sheet normal vector. As long as the sheet model shown in Fig. 10.16 is considered, other definitions of measured thickness, for example, the shortest distance between both sides of the image edges, generally give the same thickness value. We define the image edges as the zero-crossings of the second directional derivatives along the sheet normal vector, which is equivalent to the Canny edge detector [42]. Gaussian blurring is typically combined with the second directional derivatives to adjust scale as well as reduce noise.

The partial second derivative combined with Gaussian blurring for the MR image $f(\vec{x})$, for example, is given by

$$f_{xx}(\vec{x};\sigma) = g_{xx}(\vec{x};\sigma) * f(\vec{x}), \tag{10.50}$$

where

$$g_{xx}(\vec{x};\sigma) = \frac{\partial^2}{\partial x^2}\mathrm{Gauss}(\vec{x};\sigma), \tag{10.51}$$

in which $\mathrm{Gauss}(\vec{x};\sigma)$ is the isotropic 3-D Gaussian function with SD σ. The second directional derivative along $\vec{r}_{\theta,\phi}$ is represented as

$$f''(\vec{x};\sigma,\vec{r}_{\theta,\phi}) = g_{xx}(\vec{x}';\sigma) * f(\vec{x}), \tag{10.52}$$

where $\vec{x}' = R_{\theta,\phi}\vec{x}$, in which $R_{\theta,\phi}$ denotes a 3×3 matrix representing rotation which enables the normal orientation of the sheet $s_0(\vec{x};\tau)$, i.e. the x-axis, correspond to $\vec{r}_{\theta,\phi}$ (Fig. 10.16(b)). Similarly, the first directional derivative along $\vec{r}_{\theta,\phi}$ is represented as

$$f'(\vec{x};\sigma,\vec{r}_{\theta,\phi}) = g_x(\vec{x}';\sigma) * f(\vec{x}), \tag{10.53}$$

Practically, the first and second directional derivatives can be calculated in a computationally efficient manner using the Hessian matrix $\nabla^2 f(\vec{x};\sigma)$ and gradient vector $\nabla f(\vec{x};\sigma)$, respectively. The first and second directional derivatives along the normal direction $\vec{r}_{\theta,\phi}$ of the sheet structure are given by

$$f'(\vec{x};\sigma,\vec{r}_{\theta,\phi}) = \vec{r}_{\theta,\phi}^{\top}\nabla f(\vec{x};\sigma), \tag{10.54}$$

and

$$f''(\vec{x};\sigma,\vec{r}_{\theta,\phi}) = \vec{r}_{\theta,\phi}^{\top}\nabla^2 f(\vec{x};\sigma)\vec{r}_{\theta,\phi}, \tag{10.55}$$

respectively.

Thickness of sheet structures can be determined by analyzing 1-D profiles of $f''(\vec{x};\sigma,\vec{r}_{\theta,\phi})$ and $f'(\vec{x};\sigma,\vec{r}_{\theta,\phi})$ along straight line given by

$$\vec{x} = s \cdot \vec{r}_{\theta,\phi}, \tag{10.56}$$

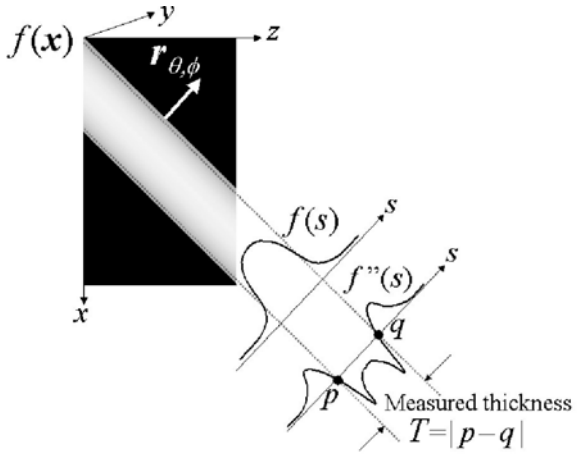
where s is a parameter representing the position on the straight line. By substituting Eq. (10.56) for $\vec{x}$ in $f''(\vec{x};\sigma,\vec{r}_{\theta,\phi})$ and $f'(\vec{x};\sigma,\vec{r}_{\theta,\phi})$,

$$f''(s) = f''(s \cdot \vec{r}_{\theta,\phi};\sigma,\vec{r}_{\theta,\phi}), \tag{10.57}$$

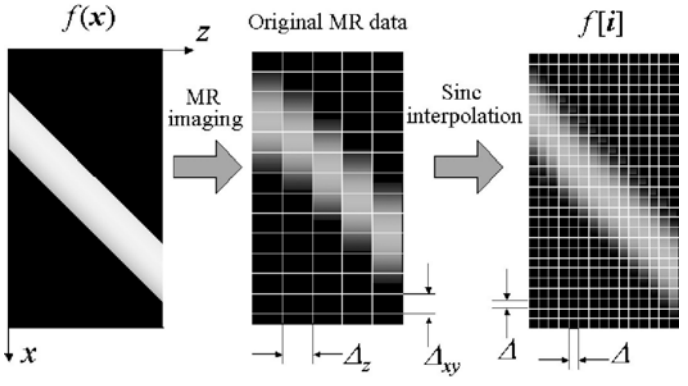
and

$$f'(s) = f'(s \cdot \vec{r}_{\theta,\phi};\sigma,\vec{r}_{\theta,\phi}), \tag{10.58}$$

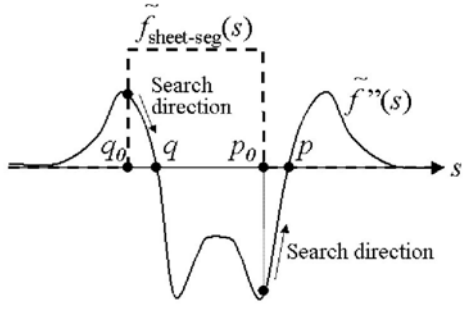
are derived respectively. Figure 10.17(a) shows a schematic diagram for the 1-D profile processing. Both sides of the boundaries for sheet structures can be



(a)



(b)



(c)

Figure 10.17: Thickness determination procedure using zero-crossings of the second directional derivatives along sheet normal direction. (a) Basic concept of thickness determination procedure. (b) Interpolation of discrete MR data. (c) Zero-crossing search procedure. (© 2004 IEEE)

defined as the points with the maximum and minimum values of $f'(s)$ among those satisfying the condition given by $f''(s) = 0$. Let $f'(s)$ have its maximum and minimum values at $s = p$ and $s = q$, respectively. The measured thickness, T is defined as the distance between the two detected boundary points, which is given by

$$T = |p - q|. \tag{10.59}$$

The procedures for thickness determination from a volume dataset in the previous section.

10.5.2 Frequency Domain Analysis of MR Imaging and Width Quantification

In order to elucidate the effects of MR imaging and postprocessing parameters, observations in the frequency domain are helpful. All the processes to obtain $f'(\vec{x}; \sigma, \vec{r}_{\theta,\phi})$ and $f''(\vec{x}; \sigma, \vec{r}_{\theta,\phi})$ from the original sheet structure $s(\vec{x}; \tau, \vec{r}_{\theta,\phi})$ are modeled as linear filtering processes excepting the magnitude operator applied in Eq. (10.49).

10.5.2.1 Modeling a Sheet Structure

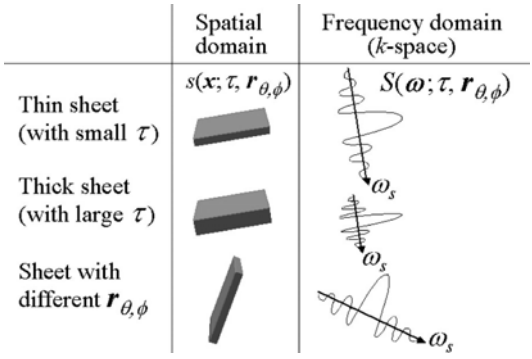
The Fourier transform of 3-D sheet structure orthogonal to the x-axis, $s_0(\vec{x}; \tau)$, is given by

$$S_0(\vec{\omega}; \tau) = \mathcal{F}\{\mathrm{Bar}(x; \tau)\} \cdot \delta(\omega_y) \cdot \delta(\omega_z), \tag{10.60}$$

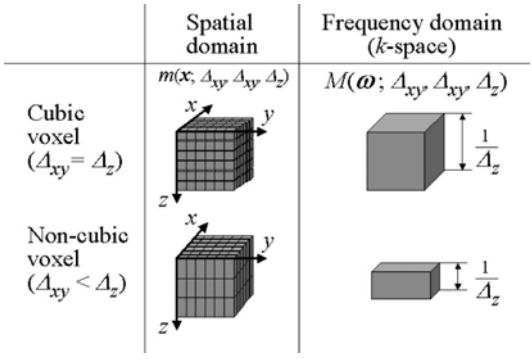
where $\mathcal{F}$ represents the Fourier transform, $\delta(\omega)$ denotes the unit impulse, and $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$. Note that $\mathcal{F}\{\mathrm{Bar}(x; \tau)\} = \tau \cdot \mathrm{Sinc}(\omega_x; \tau)$ when $L_+ = L_- = 0$ and $L_0 = 1$ in $\mathrm{Bar}(x; \tau)$. The Fourier transform of 3-D sheet structure whose normal is $\vec{r}_{\theta,\phi}$, $s(\vec{x}; \tau, \vec{r}_{\theta,\phi})$, is given by

$$S(\vec{\omega}; \tau, \vec{r}_{\theta,\phi}) = S_0(\vec{\omega}'; \tau), \tag{10.61}$$

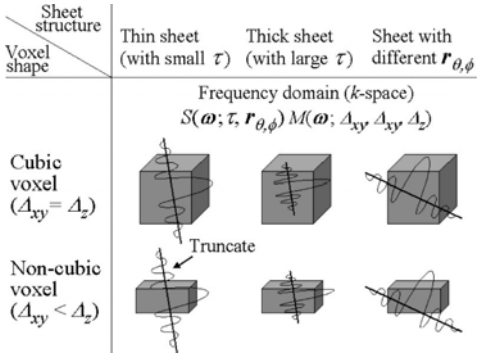
where $\vec{\omega}' = R_{\theta,\phi}\vec{\omega}$, in which $R_{\theta,\phi}$ denotes a 3×3 matrix representing rotation which enables the ω_x-axis correspond to $\vec{r}_{\theta,\phi}$ (Fig. 10.18(a)).



(a)



(b)



(c)

Figure 10.18: Frequency domain analysis of sheet structure modeling, MR imaging, and thickness determination. (a) Modeling a sheet structure. In the frequency domain, a sheet structure is basically modeled as the sinc function whose width is inversely proportional to the thickness in the spatial domain. (b) Modeling MR imaging. It is assumed here that $\Delta_{xy} = \Delta_x = \Delta_y$. The voxel size determines the frequency bandwidth of each axis, which is also inversely proportional to the size in the spatial domain. (c) Modeling MR image acquisition of a sheet structure. In the frequency domain, imaged sheet structure is essentially the band-limited sinc function. ()

In the 3-D space of the frequency domain, $S(\vec{\omega};\tau,\vec{r}_{\theta,\phi})$ has energy only in the 1-D subspace represented as a straight line given by

$$\vec{\omega} = \omega_s \cdot \vec{r}_{\theta,\phi}, \tag{10.62}$$

where ω_s is a parameter representing the position on the straight line. By substituting Eq. (10.62) for $\vec{\omega}$ in Eq. (10.61), the following is derived

$$S(\omega_s) = S(\omega_s \cdot \vec{r}_{\theta,\phi};\tau,\vec{r}_{\theta,\phi}) = \mathcal{F}\{\mathrm{Bar}(x;\tau)\}, \tag{10.63}$$

where $S(\omega_s)$ represents energy distribution along Eq. (10.62). Analysis of the degradation of 1-D distribution, $S(\omega_s)$, is sufficient to examine the effects of MR imaging and postprocessing parameters in the subsequent processes. It should be noted that $S(\omega_s)$ is the 1D sinc function when $L_- = L_+$.

10.5.2.2 Modeling MR Image Acquisition

The Fourier transform of MR PSF is given by

$$M(\vec{\omega};\Delta_x,\Delta_y,\Delta_z) = \frac{1}{\Delta_x\Delta_y\Delta_z}\mathrm{Rect}\left(\vec{\omega};\frac{1}{\Delta_x},\frac{1}{\Delta_y},\frac{1}{\Delta_z}\right). \tag{10.64}$$

where $\mathrm{Rect}(\vec{x};a_x,a_y,a_z) = \mathrm{Rect}(x;a_x)\cdot\mathrm{Rect}(y;a_y)\cdot\mathrm{Rect}(z;a_z)$ (Fig. 10.18(b)), and

$$\mathrm{Rect}(x;a) = \begin{cases} 1 & -\frac{1}{2}a \le x \le \frac{1}{2}a \\ 0 & \text{otherwise.} \end{cases} \tag{10.65}$$

By substituting Eq. (10.62) for $\vec{\omega}$ in Eq. (10.64) to obtain 1-D frequency component affecting $S(\omega_s)$, the following is derived

$$M(\omega_s) = M(\omega_s \cdot \vec{r}_{\theta,\phi};\Delta_x,\Delta_y,\Delta_z). \tag{10.66}$$

Thus, the Fourier transform of MR image of the sheet structure, $F(\omega_s)$ is given by

$$F(\omega_s) = \mathcal{F}\{|\mathcal{F}^{-1}\{S(\omega_s)M(\omega_s)\}|\}, \tag{10.67}$$

where $\mathcal{F}^{-1}$ represents the inverse Fourier transform. If $\mathcal{F}^{-1}\{S(\omega_s)M(\omega_s)\}$ is a nonnegative function, $F(\omega_s)$ is given by $F(\omega_s) = S(\omega_s)M(\omega_s)$ and all the processes can be described as linear filtering processes. Deformation of the original signal due to truncation is clearly understandable in the frequency domain (Fig. 10.18(c)).

10.5.2.3 Gaussian Derivatives of MR Imaged Sheet Structure

The Fourier transform of the second derivative of Gaussian of x is given by

$$G_{xx}(\vec{\omega};\sigma) = (\sqrt{2\pi}\sigma)^3(2\pi j\omega_x)^2\text{Gauss}\left(\vec{\omega};\frac{1}{\sqrt{2\pi}\sigma}\right), \tag{10.68}$$

and that of the second directional derivative along $\vec{r}_{\theta,\phi}$ is represented as

$$G''(\vec{\omega};\sigma,\vec{r}_{\theta,\phi}) = G_{xx}(\vec{\omega}';\sigma), \tag{10.69}$$

where $\vec{\omega}' = R_{\theta,\phi}\vec{\omega}$, in which $R_{\theta,\phi}$ denotes a 3×3 matrix representing rotation which enables the ω_x-axis correspond to $\vec{r}_{\theta,\phi}$. One-dimensional frequency component of $G''(\vec{\omega};\sigma,\vec{r}_{\theta,\phi})$ affecting $S(\omega_s)$ is given by

$$G''(\omega_s) = G''(\omega_s\cdot\vec{r}_{\theta,\phi};\sigma,\vec{r}_{\theta,\phi}). \tag{10.70}$$

Similarly, 1-D component of the first directional derivative of Gaussian, $G'(\omega_s)$, is obtained.

Finally the Fourier transforms of $f''(s)$ and $f'(s)$ are derived and given by

$$F''(\omega_s) = F(\omega_s)G''(\omega_s) \tag{10.71}$$

and

$$F'(\omega_s) = F(\omega_s)G'(\omega_s), \tag{10.72}$$

respectively.

The 1-D profiles along the sheet normal direction of the Gaussian derivatives of MR imaged sheet structures (Eqs. (10.57) and (10.58)) are obtained by inverse Fourier transform of Eqs. (10.71) and (10.72), and then thickness is determined according to the procedure shown in Fig. 10.17(a). While simulating the MR imaging and Gaussian derivative computation described in section 10.5.1 essentially requires 3-D convolution in the spatial domain, only 1-D computation is necessary in the frequency domain, which drastically reduces computational cost. In the following sections, we examine the effects of various parameters, which are involved in the sheet model, MR imaging resolution, and thickness determination processes, on measurement accuracy. Efficient computational methods of simulating MR imaging and postprocessing thickness determination processes are essential, and thus simulating the processes by 1-D signal processing in the frequency domain is regarded as the key to comprehensive analysis.

10.5.3 Accuracy Evaluation by Numerical Simulation

In order to examine the effects of various parameters on the accuracy of thickness determination, numerical simulation based on the theory described in the previous section was performed. The parameters used in the simulation were classified into the following three categories: τ, $\vec{r}_{\theta,\phi}$, L_-, L_0, and L_+ for defining sheet structures; Δ_x, Δ_y, and Δ_z for determining MR imaging resolution; and Gaussian SD, σ, used in computer postprocessing for thickness determination.

We assumed that the estimated sheet thickness T was obtained under the condition that the sheet normal orientation $\vec{r}_{\theta,\phi}$ was known. The numerical simulation was performed in the frequency domain exactly in the same manner as described in section 10.5.2. Based on sheet structure parameters τ, $\vec{r}_{\theta,\phi}$, L_-, L_0, and L_+, MR imaging parameters Δ_x, Δ_y, and Δ_z, and postprocessing parameters σ, $F'(\omega_s)$ and $F''(\omega_s)$ were obtained by 1-D computation in the frequency domain according to Eqs. (10.71) and (10.72), respectively. And then, $f'(s)$ and $f''(s)$ were obtained by inverse Fourier transform of $F'(\omega_s)$ and $F''(\omega_s)$, respectively. Using $f'(s)$ and $f''(s)$, thickness T were estimated using Eq. (10.59). Finally, estimated thickness T was compared with the actual thickness τ to reveal the limits on accuracy. It should be noted that only 1-D computation was necessary for 3D thickness determination in our numerical simulation.

In the simulation, the effect of anisotropic resolution of MR volume data was the focus. Let $\Delta_{xy}(= \Delta_x = \Delta_y)$ be the pixel size within the slices. Resolution of MR volume data is typically anisotropic because they usually have lower resolution along the third direction (orthogonal to the slice plane) than within slices. Hence, it can be assumed that the resolution along the z-axis is lower than that in the xy-plane and that pixels in the xy-plane are square, i.e. $\Delta_{xy} \leq \Delta_z$, and a measure of voxel anisotropy can be defined as $\frac{\Delta_z}{\Delta_{xy}}$. In the simulations, we assumed that

$$\Delta_{xy} = 1, \tag{10.73}$$

without loss of generality, and thus,

$$\text{voxel anisotropy} = \frac{\Delta_z}{\Delta_{xy}} = \frac{\Delta_z}{1} = \Delta_z \geq 1. \tag{10.74}$$

We performed the above described numerical simulation with different combinations of τ, $\vec{r}_{\theta,\phi}$, L_-, L_0, L_+, Δ_z, and σ.

Table 10.3: Parameter values used in numerical simulations

	Sheet structure			MR imaging		
	Thickness	Orientation		Voxel size		Postprocessing Gaussian SD
Figure	τ	θ	ϕ	Δ_{xy}	Δ_z	σ
10.19	Variable	0°, 45°	0°	1	1	$\frac{1}{2}$, $\frac{\sqrt{2}}{2}$, 1
10.20(a)	1, 2, 3, 4, 5, 6	Variable	0°	1	1, 2, 4	$\frac{\sqrt{2}}{2}$
10.20(b)	2	Variable	0°	1	Variable	$\frac{\sqrt{2}}{2}$
10.20(c)	1, 2, 3, 4, 5, 6	Variable	0°	1	2	Anisotropic $\sigma_{xy} = \frac{\sqrt{2}}{2}$, $\sigma_z = \frac{\sqrt{2}}{2}\Delta_z$

The unit of dimension for the following simulation results was Δ_{xy}, i.e., $\Delta_{xy} = 1$ was assumed as described in the previous section. Thus, other parameters (τ, T, Δ_z, σ) were normalized by Δ_{xy}, and voxel anisotropy was represented as $\Delta_z = (\frac{\Delta_z}{\Delta_{xy}} = \frac{\Delta_z}{1})$. In the simulation, we used $L_0 = 200$ and $L_- = L_+ = 100$ for the bar profile. These parameter values were determined so that the bar profile was symmetric and the magnitude operator in Eq. (10.49) did not affect the results. Table 10.3 summarizes the parameter values used in the numerical simulations described below.

10.5.3.1 Effects of Gaussian Standard Deviation in Postprocessing

Figure 10.19 shows the effects of the standardg deviation (SD), σ, in Gaussian blurring. In Fig. 10.19(a), the relations between true thickness τ and measured thickness T are shown for three σ values ($\frac{1}{2}$, $\frac{\sqrt{2}}{2}$, 1) when Δ_z is equal to 1, i.e. in the case of isotropic voxel. The relation is regarded as ideal when $T = \tau$, which is the diagonal in the plots of Fig. 10.19(a). For each σ value, the relations were plotted using two values of sheet normal orientation θ (0°, 45°), while ϕ was fixed to 0°. Strictly speaking, voxel shape is not perfectly isotropic even when Δ_z is equal to 1 because the shape is not spherical. Thus, slight dependence on θ was observed.

In order to observe the deviation from $T = \tau$ more clearly, we defined the error as $E = T - \tau$. Figure 10.19(b) shows the plots of error E instead of T. With $\sigma = \frac{1}{2}$, considerable ringing was observed for error E. With $\sigma = 1$, error magnitude $|E|$ was significantly large for small τ (around $\tau = 2$). With $\sigma = \frac{\sqrt{2}}{2}$,

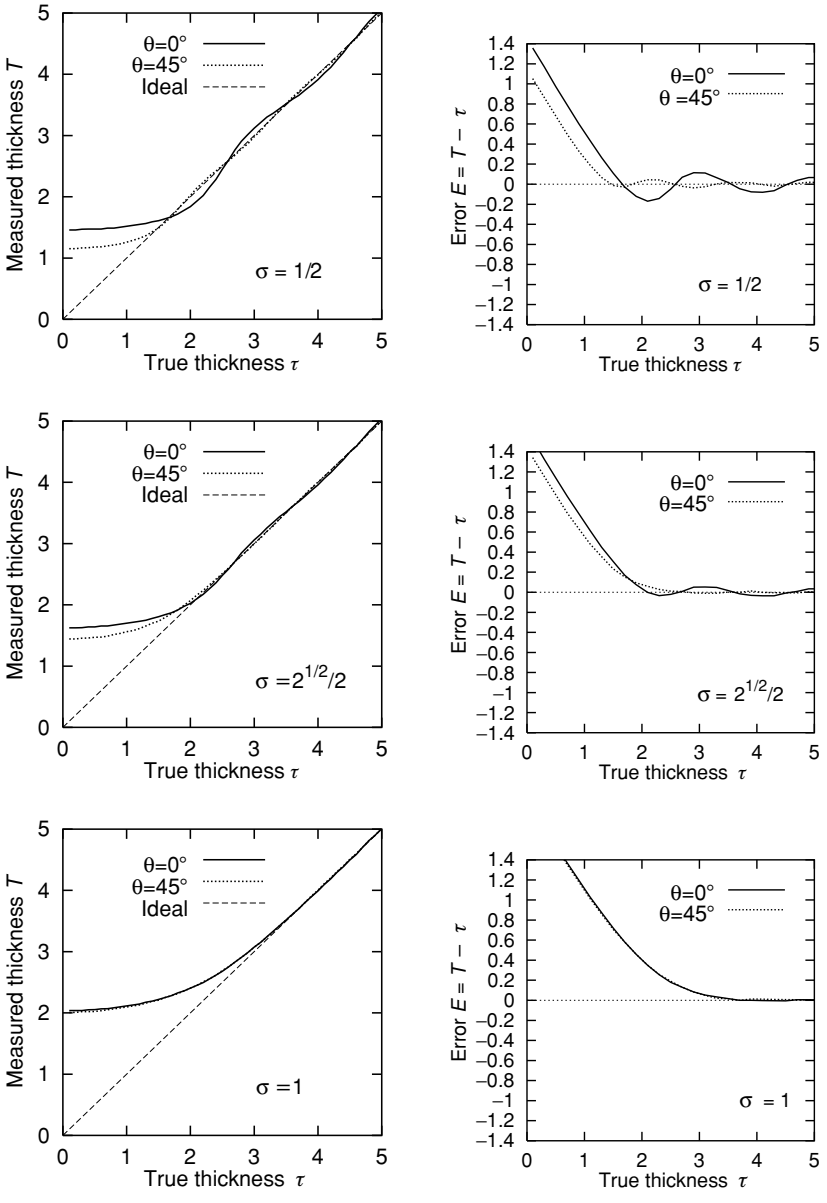


Figure 10.19: Effects of Gaussian SD, σ in postprocessing for thickness determination with isotropic voxel. The unit is Δ_{xy}. $\phi = 0°$. (a) Relations between true thickness τ and measured thickness T. (b) Relations between true thickness τ and error $T - \tau$. (© 2004 IEEE)

however, ringing became small and error magnitude $|E|$ was sufficiently small around $\tau = 2$. $\sigma = \frac{\sqrt{2}}{2}$ gave a good compromise optimizing the trade-off between reducing the ringing and improving the accuracy for small τ. Actually, error magnitude $|E|$ is guaranteed to satisfy $|E| < 0.1$ for $\tau > 2.0$ with $\sigma = \frac{\sqrt{2}}{2}$, while $|E| < 0.1$ for $\tau > 3.2$ with $\sigma = \frac{1}{2}$ and, $|E| < 0.1$ for $\tau > 2.9$ with $\sigma = 1$. Based on this result, we used $\sigma = \frac{\sqrt{2}}{2}$ in the following experiments if not specified.

10.5.3.2 Effects of Voxel Anisotropy in MR Imaging

Figure 10.20(a) shows the effects of sheet normal orientation θ and voxel anisotropy Δ_z on measured thickness T. The relations between measured thickness T and sheet normal orientation θ for six values of true thickness τ (1, 2, 3, 4, 5, 6) were plotted when three different values of voxel anisotropy Δ_z (1, 2, 4) were used. The relations were regarded as ideal when $T = \tau$ for any θ, which is the horizontal in the plots of Fig. 10.20(a). When $\Delta_z = 1$, the relations were highly close to the ideal for $\tau > 2$. When $\Delta_z = 2$ and $\Delta_z = 4$, significant deviations from the ideal were observed for $\theta > 15^\circ$ and $\theta > 30^\circ$, respectively.

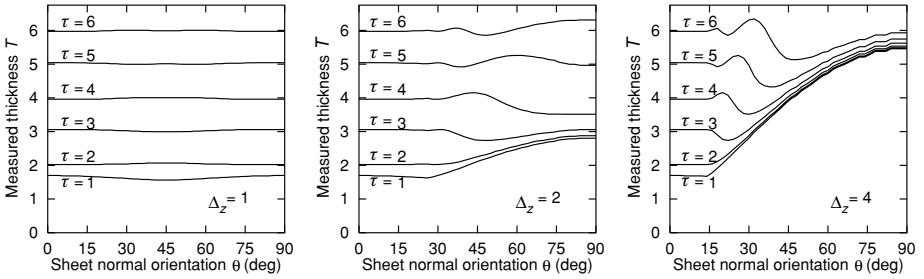
Figure 10.20(b) shows the plots of the maximum θ at which error magnitude $|E|$ is guaranteed to satisfy $|E| < 0.1$, $|E| < 0.2$, and $|E| < 0.4$ for $\tau = 2$ with varied voxel anisotropy Δ_z. These plots clarify the range of θ where the deviation from the ideal is sufficiently small. There was no significant difference between the plots for $\tau = 2$ and different values of τ (for $\tau > 2$).

10.5.3.3 Using Anisotropic Gaussian Blurring Based on Voxel Anisotropy

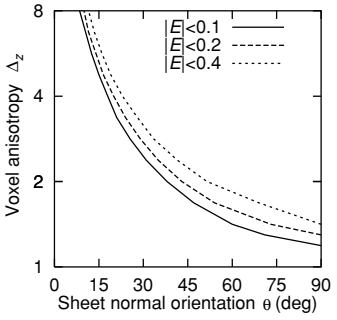
We have assumed that Gaussian blurring combined with derivative computation is isotropic as shown in Eq. (10.51). Another choice is to use anisotropic Gaussian blurring corresponding to voxel anisotropy, which is given by

$$g_{xx}(\vec{x}; \sigma_{xy}, \sigma_z) = \frac{\partial^2}{\partial x^2}\mathrm{Gauss}(x, y; \sigma_{xy})\mathrm{Gauss}(z; \sigma_z), \tag{10.75}$$

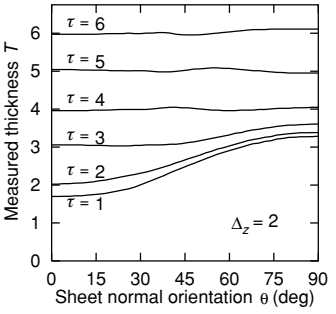
where σ_z and σ_{xy} are determined so as to satisfy $\frac{\sigma_z}{\sigma_{xy}} = \frac{\Delta_z}{\Delta_{xy}}$, and thus $\sigma_z = \Delta_z \sigma_{xy}$ because we assumed $\Delta_{xy} = 1$. Figure 10.20(c) shows plots of measured thickness obtained using anisotropic Gaussian blurring when $\Delta_z = 2$ and $\sigma_{xy} = \frac{\sqrt{2}}{2}$. The plots using anisotropic Gaussian blurring were closer to the ideal for $\tau \geq 4$ and any θ, while those using isotropic one were closer for $\tau \geq 2$ and $\theta < 30^\circ$.



(a)



(b)



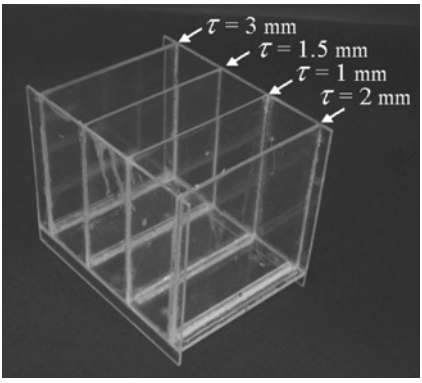
(c)

Figure 10.20: Effects of voxel anisotropy Δ_z in MR imaging and anisotropic Gaussian blurring on measured thickness T. The unit is Δ_{xy}. $\sigma = \frac{2^{\frac{1}{2}}}{2}$ and $\phi = 0^\circ$. (a) Relations between measured thickness T and sheet normal orientation θ with different τ values. (b) Plots of maximum θ at which error magnitude $|E|$ is guaranteed to satisfy $|E| < 0.1$, $|E| < 0.2$, and $|E| < 0.4$ for $\tau = 2$ while voxel anisotropy Δ_z is varied (where $E = T - \tau$). (c) Relations between true thickness τ and measured thickness T with the use of anisotropic Gaussian blurring based on voxel anisotropy. $\sigma_{xy} = \frac{2^{\frac{1}{2}}}{2}$ and $\sigma_z = \frac{2^{\frac{1}{2}}}{2}\Delta_z$. (© 2004 IEEE)

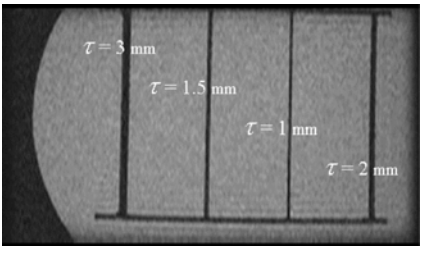
10.5.4 Validating the Numerical Simulation by *in Vitro* Experiments

To validate the numerical simulation, the postprocessing method for thickness determination was used to measure real MR images of two different objects, an acrylic plate phantom.

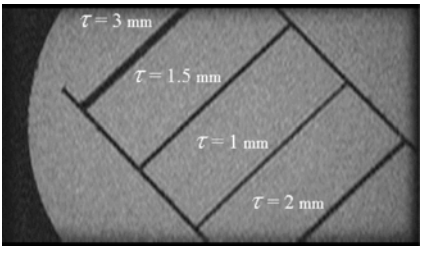
A phantom of sheet-like objects with known thickness was used. It consisted of four acrylic plates of 80×80 (mm^2) with thickness $\tau = 1.0, 1.5, 2.0$, and 3.0 (mm), placed parallel to each other with an interval of 30 mm (Fig. 10.21(a)).



(a)



$\theta = 0^\circ$ and $\phi = 0^\circ$



$\theta = 45^\circ$ and $\phi = 0^\circ$

(b)

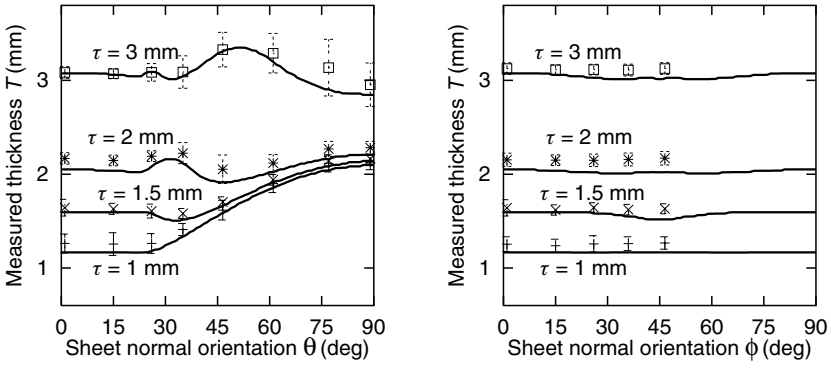
Figure 10.21: Acrylic plate phantom and its MR images. (a) Physical appearance. (b) MR images. The horizontal and vertical axes of the images correspond to the x-axis and z-axis, respectively. The voxel size was $\Delta_{xy} = 0.625$ mm and $\Delta_z = 1.5$ mm. As can be easily observed by naked eye, the acrylic plate with $\tau = 1$ mm appears to be imaged slightly thicker in $\theta = 45^\circ$ and $\phi = 0^\circ$ than $\theta = 0^\circ$ and $\phi = 0^\circ$. (© 2004 IEEE)

The phantom was submerged in a water bath so that the background (water) showed higher intensity as contrasted to low intensity objects (acrylic plates). Three-dimensional MR images (TR/TE/flip angle/matrix/FOV/slice thickness: 12.8 ms/5.6 ms/5/256×256/160 mm/1.5 mm) of the phantom were obtained using a fast spoiled gradient-echo sequence (FSPGR). The voxel size was $\Delta_{xy} = 0.625 (= \frac{160}{256})$ (mm) and $\Delta_z = 1.5$ (mm). Thus,

$$\text{voxel anisotropy} = \frac{\Delta_z}{\Delta_{xy}} = \frac{1.5}{0.625} = 2.4. \tag{10.76}$$

Thirteen datasets of 3D MR images were acquired with different normal positions of the phantom plates, eight with variable θ ($\theta = 0$, 15, 25, 35, 45, 60, 75, and 90 degrees) and fixed ϕ ($\phi = 0$), and five with variable ϕ ($\phi = 0$, 15, 25, 35, and 45 degrees) and fixed θ ($\theta = 0$). In the obtained MR images, we observed $L_- = L_+ = 40$ and $L_0 = 0$. Figure 10.21(b) shows examples of the MR images. We compared actually measured thickness from the real MR data with the computational thickness calculated by the numerical simulations.

Figure 10.22 shows the averages and the SDs of the actually measured (*in vitro*) thickness from the MR data of the phantom imaged with different



(a) (b)

Figure 10.22: Comparison of simulated thickness and *in vitro* thickness determined from MR images of acrylic plate phantom. $\Delta_{xy} = 0.625$ mm, $\Delta_z = 1.5$ mm, and $\sigma = \frac{2^{\frac{1}{2}}}{2}\Delta_{xy}$. For *in vitro* thickness, its average and SD values are indicated by symbols and error bars. (a) Dependences on sheet normal orientation θ. (b) Dependences on sheet normal orientation ϕ. Note that the dependence on ϕ is theoretically equivalent to the dependence on θ when the anisotropy is $\frac{\Delta_z}{\Delta_{xy}} = 1$. (© 2004 IEEE)

θ and ϕ and the plots of the simulated thickness representing the dependences on sheet normal orientation θ and ϕ. Figures 10.22(a) and 10.22(b) show the plots of the dependences of θ and ϕ with $\sigma = \frac{\sqrt{2}}{2}\Delta_{xy}$, respectively. Good agreement between the simulated and the *in vitro* thicknesses was observed in both cases although the *in vitro* thicknesses was slightly greater than the simulated thickness. The biases, i.e., the difference between the simulated thickness and the average of *in vitro* thickness, were predominantly around 0.1 mm or less (except for $\theta = 75^\circ$ of $\tau = 3$ mm), and the SDs of the *in vitro* thickness were mostly within 0.1 mm (except for $\theta = 45^\circ$ of $\tau = 2$ mm and $\theta \geq 35^\circ$ of $\tau = 3$ mm). It should be noted that the dependence on ϕ is theoretically equivalent to the dependence on θ when the anisotropy is $\frac{\Delta_z}{\Delta_{xy}} = 1$.

10.6 Concluding Remarks

We have presented a framework for multiscale analysis of the second-order local structures in medical volume data based on the analysis of the Hessian matrix. Multiscale filtering methods for enhancement of sheet, line, and blob structures were formulated. The guidelines for the filter design were clarified based on detailed analyses of single- and multiscale filter responses using mathematical local structure models. Further, formal approaches to description and quantification of sheet and line structures were presented. The accuracy of width quantification of sheet structures were theoretically analyzed and its inherent limits due to imaging resolution and postprocessing parameters were derived. In this chapter, we purely focus on *local* structures. Future work will include grouping these local structures to obtain higher-level descriptions incorporating global structures.

10.7 Acknowledgments

The author thanks Dr. Ron Kikinis and Dr. Shin Nakajima at Harvard Medical School and Brigham and Women's Hospital for providing MR data of a brain, Dr. Hironobu Ohmatsu of the National Cancer Center, Japan, for providing CT data of a chest, Dr. Nobuyuki Shiraga at Keio University for providing abdominal CT data, Dr. Shigeyuki Yoshida at Osaka University for providing a CT data of a

chest, and Dr. Katsuyuki Nakanishi, Dr. Hisashi Tanaka, Dr. Nobuhiko Sugano, and Dr. Takashi Nishii at Osaka University for providing hip joint MR data and phantom MR data. The author also thanks all the above researchers and Prof. Shinicni Tamura at Osaka University for fruitful discussion.

Questions

1. *Summarize a series of procedures for multiscale enhancement filtering described in this chapter from input original images to output final filter-enhanced images.*

2. *Explain the parameters involved in the procedures and discuss how to select these parameters.*

3. *Derive the mathematical formula of the width response curves shown in Fig. 10.3.*

4. *Discuss the effect of the anisotropic resolution (voxel shape) of input volume data on multiscale enhancement filtering.*

Bibliography

[1] Knutsson, H., Representing local structure using tensors, In: Proceedings of 6th Scandinavian Conference on Image Analysis, 1989, pp. 244–251.

[2] Westin, C.-F., A Tensor Framework for Multidimensional Signal Processings, Ph.D. Thesis No. 348, Linköping University, Sweden, 1994.

[3] Koller, T . M., Gerig, G., Szekely, G., and Dettwiler, D., Multiscale detection of curvilinear structures in 2-D and 3-D image data, In: Proceedings of Fifth International Conference on Computer Vision, 1995, pp. 864–869.

[4] Aylward, S., Bullitt, E., Pizer, S., and Eberly, D., Intensity ridge and widths for tubular object segmentation and description, In: Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 1996, pp. 131–138.

[5] Sato, Y., Nakajima, S., Atsumi, H., Koller, T., Gerig, G., Yoshida, S., and Kikinis, K., 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images, In: Lecture Notes in Computer Science, Vol. 1205, pp. 213–222, 1997. Proceedings of CVRMed-MRCAS'97, Glenoble, France.

[6] Lorenz, C., Carlsen, I.-C., Buzug, T. M., Fassnacht, C., and Wesse, J., Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images, In: Lecture Notes in Computer Science, Vol. 1205, pp. 233–242, 1997. Proceedings of CVRMed-MRCAS'97, Glenoble, France.

[7] Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Guido, G., and Kikinis, R., Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images, Med. Image Anal., Vol. 2, No. 2, pp. 143–168, 1998.

[8] Frangi, A., Niessen, W., Vincken, K., and Viergever, M., Multiscale vessel enhancement filtering, Vol. 1426 In: Proceedings of MICCAI'98, Boston, Massachusetts, 1998, pp. 130–137.

[9] Westin, C.-F., Warfield, S., Bhalerao, A., Mui, L., Richolt, J., and Kikinis, R., Tensor controlled local structure enhancement of CT images for bone segmentation, In: Lecture Notes in Computer Science, Vol. 1426, pp. 1205–1212, 1998. Proceedings of MICCAI'98, Boston, Massachusetts.

[10] Sato, Y., Kubota T., Nakanishi K., Sugano N., Nishii T., Ohzono K., Nakamura H., Ochi O., and Tamura S., Three-dimensional reconstruction and quantification of hip joint cartilages from magnetic resonance images, In: Lecture Notes in Computer Science, Vol. 1679, pp. 338–347, 1999. Proceedings of MICCAI'99, Cambridge, UK.

[11] Sato, Y., Westin, C.-F., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., and Kikinis, R., Tissue classification based on 3D local intensity structures for volume rendering, IEEE Trans. Visual. Comput. Graphics, Vol. 6, No. 2, pp. 160–180, 2000.

[12] Sato, Y. and Tamura S., Detection and quantification of line and sheet structures in 3-D images, In: Lecture Notes in Computer Sceinece, Vol. 1935, pp. 164–165, 2000. Proceedings of MICCAI 2000, Pittsburgh, Pennsylvania.

[13] Aylward, S. R. and Bullitt, E., Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction, IEEE Trans. Med. Imaging, Vol. 21, No. 2, pp. 61–75, 2002.

[14] Suri, J. S., Liu, K., Reden L., and Laxminarayan, S. N., White and black blood volumetric angiographic filtering: Ellipsoidal scale-space approach, IEEE Trans. Inform. Tech. Biomed., Vol. 6, No. 2, pp. 142–158, 2002.

[15] Suri, J. S., Liu, K., Reden L., and Laxminarayan, S. N., A review on MR vascular image processing algorithms: Acquisition and prefiltering, Part I, IEEE Trans. Inform. Tech. Biomed., Vol. 6, No. 4, pp. 324–337, 2002.

[16] Suri, J. S., Liu, K., Reden L., and Laxminarayan, S. N., A review on MR vascular image processing algorithms: Skeleton versus nonskeleton approaches, Part I, IEEE Trans. Inform. Tech. Biomed., Vol. 6, No. 4, pp. 338–350, 2002.

[17] Sato, Y., Nakanishi, K., Tanaka, H., Nishii, T., Sugano, N., Nakamura, H., Ochi, T., and Tamura, S., Limits to the accuracy of 3D thickness measurement in magnetic resonance images, In: Lecture Notes in Computer Science, Vol. 2208, pp. 803–810, 2001. Proceedings of MICCAI2001, Utrecht, The Netherlands.

[18] Sato, Y., Tanaka, H., Nishii, T., Nakanishi, K., Sugano, N., Kubota, T., Nakamura, H., Yoshikawa, H., Ochi, T., and Tamura, S., Limits on the accuracy of 3D thickness measurement in magnetic resonance images—Effects of voxel anisotropy, IEEE Trans. Med. Imaging, Vol. 22, No. 9, pp. 1076–1088, 2003.

[19] Haralick, R. M., Watson, L. T., and Laffey, T. J., The topographic primal sketch, Int. J. Robotic Res., Vol. 2, No. 1, pp. 50–72, 1983.

[20] Marr, D., Vision—A Computational Investigation into the Human Representation and Processing of Visual Information, W. H. Freeman, New York, 1982.

[21] Lindeberg, T., On scale selection for differential operators, Proc. 8th Scandinavian Conference on Image Analysis, pp. 857–866, 1993.

[22] Lindeberg, T., Feature detection with automatic scale selection, Int. J. Comput. Vision, Vol. 30, No. 2, pp. 77–116, 1998.

[23] Lindeberg, T., Edge Detection and ridge detection with Automatic Scale Selection, Int. J. Comput. Vision, Vol. 30, No. 2, pp. 117–154, 1998.

[24] Hylton, N. M., Simovsky, I., Li, A . J., and Hale, J . D., Impact of section doubling on MR angiography, Radiology, Vol. 185, No. 3, pp. 899–902, 1992.

[25] Du, Y. P., Parker, D. L., Davis, W. L., and Cao, G., Reduction of partial-volume artifacts with zero-filled interpolation in three-dimensional MR angiography, J. Magn. Reson. Imaging, Vol. 4, No. 5, pp. 733–741, 1995.

[26] Kikinis, R., Gleason, P. L., Moriarty, T. M., Moore, M. R., Alexander, E., III, Stieg, P. E., Matsumae, M., Lorensen, W. E., Cline, H. E., Black, P. M., Jolesz, F. A., Computer-assisted interactive three-dimensional planning

for neurosurgical procedures, Neurosurgery, Vol. 38, No. 4, pp. 640–651, 1996.

[27] Nakajima, S., Atsumi, H., Kikinis, R., Moriarty, T. M., Metcalf, D. C., Jolesz, F. A., and Black, P. M., Use of cortical surface vessel registration for image-guided neurosurgery, Neurosurgery, Vol. 40, No. 6, pp. 1201–1210, 1997.

[28] Levoy, M., Display of surfaces from volume data, IEEE Comput. Graphics Appl., Vol. 8, No. 3, pp. 29–37, 1988.

[29] Lacroute, P. and Levoy M., Fast volume rendering using a shear-warp factorization of the viewing transform, In: Proceedings of SIGGRAPH'94, 1994, pp. 451–458.

[30] Vandarbrug, S. J., Semilinear line detectors, Comput. Graphics Image Process., Vol. 4, pp. 287–293, 1975.

[31] Burt, P. J. and Adelson E. H., The Laplacian pyramid as a compact image code, IEEE Trans. Commun., Vol. 31, No. 4, pp. 532–540, 1983.

[32] Shiraga, N., Sato, Y., Kohda, E., Okada, Y., Sato, K., Hasebe, T., Hiramatsu, K., Kikinis, R., and Jolesz, F. A., Three dimensional display of the osteosclerotic lesion by volume rendering method, Nippon Acta Radiol. Vol. 58, No. 2, p. S84, 1998.

[33] Shimizu, A., Hasegawa, J., and Toriwaki J., Minimum directional difference filter for extraction of circumscribed shadows in chest X-ray images and its characteristics, IEICE Trans., Vol. J-76D-II, No. 2, pp. 241–249, 1993.

[34] Giger, M. L., Bae, K. T., and MacMahon, H., Computerized detection of pulmonary nodules in computed tomography images, Invest. Radiol., Vol. 24, No. 4, pp. 459–465, 1994.

[35] Kanazawa, K., Kubo, M., Niki, N., Satoh, H., Ohmatsu, H., Eguchi, K., and Moriyama, N., Computer aided screening system for lung cancer based on helical CT images, In: Lecture Notes in Computer Science, Vol. 1131, pp. 223–228, 1996. Proceedings of Visualization in Biomedical Computing, Hamburg, Germany.

[36] Steger, C., An unbiased detector of curvilinear structures, IEEE Trans. Patt. Anal. Machine Intell., Vol. 20, No. 2, pp. 113–125, 1998.

[37] Nakanishi, N., Tanaka, H., Nishii, T., Masuhara, K., Narumi, Y., and Nakamura, H., MR evaluation of the articular cartilage of the femoral head during traction, Acta Radiol., Vol. 40, No. 1, pp. 60–63, 1999.

[38] Nakanishi, K., Tanaka, H., Sugano, N., Sato, Y., Ueguchi, T., Kubota, T., Tamura, S., and Nakamura, H., MR-based three-dimensional presentation of cartilage thickness in the femoral head, Euro. Radiol., Vol. 11, No. 11, pp. 2178–2183, 2001.

[39] Parker, D. L., Du, Y. P., and Davis, W. L., The voxel sensitivity function in Fourier transform imaging: applications to magnetic resonance angiography, Magn. Reson. Med., Vol. 33, No. 2, pp. 156–162, 1995.

[40] Hoogeveen, R. M., Bakker, C. J. G., and Viergever, M. A., Limits to the accuracy of vessel diameter measurement in MR angiography, J. Magn. Reson. Imaging, Vol. 8, No. 6, pp. 1228–1235, 1998.

[41] Steckner, M. C., Drost, D. J., and Prato, F. S., Computing the modulation transfer function of a magnetic resonance imager, Med. Phy., Vol. 21, No. 3, pp. 483–489, 1994.

[42] Canny, J., A computational approach to edge detection, IEEE Trans. Patt. Anal. Machine Intell., Vol. 8, No. 6, pp. 679–698, 1986.

Chapter 11

A Knowledge-Based Scheme for Digital Mammography

Sameer Singh[1] *and Keir Bovis*[2]

11.1 Introduction

The automated detection of lesions in the breast is important. The area of computer-aided detection (CAD) in digital mammography is devoted to developing sophisticated image analysis tools that can automatically detect breast lesions. The whole process can be viewed as a pipeline of subprocesses that are aimed at finding regions of interest (ROI) and classifying them in breast images. These processes (layers) are common to most medical imaging applications and involve image preprocessing, enhancement, segmentation, feature extraction, classification, and postprocessing for reducing false positives. There is a variety of algorithms for these processes available in medical imaging literature but little to guide their selection. There are only a few comparative studies that exhaustively compare different algorithms on large datasets and correlate the success of the algorithm with the type of data used. Most clinical studies use a preselected set of image analysis algorithms that are uniformly applied to all images. In our opinion, this practice is not good. In this chapter we demonstrate the use of a knowledge-based framework that integrates the various layers of analysis under an adaptive scheme. The main emphasis is to have at our disposal more than one algorithm per layer to produce the same type of output, and then

[1] Pann Research, Department of Computer Science, University of Exeter, Exeter EX4 4QF, UK
[2] Met Office, Fitzroy Road, Exeter EX1 3PB, UK

based on the properties of the image under consideration, predict the single best algorithm to be applied at each layer from this set. We demonstrate that this scheme of work has significant advantages over a nonadaptive structure (where only one algorithm is available per layer and it is fixed for all images in the dataset).

We aim to answer the following questions: (a) What is a knowledge-based framework? We discuss the components of this framework in section 11.2 putting it in the context of previous research. (b) How does the image enhancement layer work in this framework? This is detailed in section 11.3 where we discuss measures of image viewability based on enhancement, and demonstrate the role of good enhancement in image segmentation. We also propose two new mapping schemes that can map the image features to chosen enhancement methods. (c) How does the image segmentation layer work within the knowledge-based framework? In section 11.4 we detail the implementation of sophisticated Gaussian mixture models in both supervised and unsupervised modes, with an expert combination framework and compare them on overlap measures. (d) What are the different strategies for reducing false positives? In section 11.5 we discuss several postprocessing steps that are aimed at reducing the number of false positives per image. (e) Is the adaptive knowledge-based framework superior to a nonadaptive scheme that uses the same algorithms across all images uniformly? We discuss our results on this issue in section 11.6 where we show the relative superiority of the adaptive framework.

11.2 Knowledge-Based Framework

The CAD scheme detailed in this chapter is based on an adaptive framework. An adaptive framework is capable of modifying itself such that it is more suitable to the environment within which it operates. Within the context of CAD, an adaptable component or a framework, attempts to automatically optimize the lesion detection process for a given mammogram. Broadly speaking, an adaptive characteristic can be built into CAD in three different ways: (1) Using a deterministic component; (2) knowledge-based component; (3) with a knowledge-based framework. Each approach may be used in combination with the others. These approaches are summarized below.

1. *Deterministic component*: This is the most common strategy for introducing adaptability into a CAD scheme. Typically the component method is fixed but parameters are adaptively determined on a per-image basis. The parameter setting is either performed in a deterministic manner, based directly on an observed feature of the image, e.g. variance of gray scales, or empirically through experimentation. In the past, this approach has been applied to each of the three CAD components, e.g., adaptive contrast enhancement methods to perform an optimal contrast enhancement on an image based on local neighborhoods [1, 2], adaptive segmentation techniques utilizing adaptive clustering [3–5], or thresholding techniques [6–8] to segment an image and adaptive classification methods in the reduction of false-positive regions.

2. *Knowledge-based component*: An alternative strategy for setting a given component's parameters is to learn the optimal parameter settings for an individual or group of images using machine learning techniques. The mapping between the parameter settings and images is achieved using a global image characteristics.

3. *Knowledge-based framework*: An extension to the knowledge-based component is to use machine learning principles to learn the utility of a particular component technique for an individual or group of images. Such a knowledge-based framework would be capable of drawing on a variety of different techniques to meet the objectives of each CAD component. The framework would support the definition of an optimal pipeline through the CAD pyramid. To date, no research has been presented into the use of knowledge-based frameworks in medical imaging CAD schemes. We now highlight some past research into knowledge-based components used in CAD schemes to put our proposed model in context.

11.2.1 A Review of Knowledge-Based Components in Medical Imaging CAD Schemes

Table 11.1 summarizes medical imaging CAD studies that have used a knowledge-based component. The studies are categorized according to their approach to representing the knowledge within the component. Each of the approaches is described in more detail below.

Table 11.1: Summary of studies utilizing knowledge-based components, grouped according to their approach

		Methodology			
Study	*Modality*	*part*	*ann*	*usr*	*cbr*
Zheng *et al.* [9]	Mammography (x-ray)	√			
Matsubara *et al.* [10]	Mammography (x-ray)	√			
Lai and Fang [11, 12]	Angiography		√		
Pitiot *et al.* [13]	MRI		√		
Sha and Sutton [14]	MRI		√		
Fenster and Kender [15]	CT			√	
Perner [16]	CT				√

part = image grouping; *ann* = multistage neural networks; *usr* = user interaction; *cbr* = case-based reasoning.

11.2.1.1 Knowledge Representation by Image Grouping on Various Criteria

This approach to implementing a knowledge-based component attempts to adaptively determine optimum parameter settings for groups of images on the basis of image feature vectors. The feature vector is used to group images accordingly, that in turn serve as a form of *a priori* knowledge for use in subsequent components. In this way components may be trained to operate on particular image groupings with different parameter settings.

In mammography, Zheng *et al.* [9] propose an adaptive computer-aided diagnosis scheme optimized on the basis of the characterization of the mammogram. The rule-based system proposes a difficulty index (DI). This is computed as the weighted sum of nine histogram-based features calculated from a separate training set. The computed DI score is used in conjunction with a banding scheme, based on empirically determined values corresponding to easy, moderately difficult and difficult groupings following human interpretation. An expert radiologist evaluates each mammogram and determines the group boundaries. The authors propose the use of a rule-based classification scheme such that different classification rules are independently set for the three different difficulty groups in training. On a locally defined dataset of 428 digitized mammograms (abnormal $n = 220$, normal $n = 208$), the authors report the simple adaptive

scheme reduced the average number of false-positive detections from 0.85 to 0.53 per image.

Matsubara *et al.* [10] proposed the use of an image grouping scheme for digitzed mammograms. In their study, images are assigned to one of four categories based on histogram analysis of the image gray scales. Subsequent image-processing operations, such as threshold-based segmentation and region classification operate on parameters defined empirically and independently within each category. The authors use this scheme to ignore high-density mammograms. On a small dataset of 30 images, the authors report a sensitivity of 93%.

11.2.1.2 Knowledge Representation with Multistage Neural Networks

An alternative to a hard bounded grouping scheme such as those proposed above is in the use of soft decision boundaries. These knowledge-based components utilize a mixture-of-experts paradigm. Lai and Fang [11, 12] proposed the use of a hierarchical neural network to model the optical transformation of a 12-bit magnetic resonance image (MRI) into an 8-bit representation for display on a computer monitor. This optimal optical transformation is crucial if the expert radiologist is to effectively interpret the displayed image. The authors cite a major obstacle to the implementation of a simple linear solution as the difference in optimal parameters for different types of MR images. For example, 3D angiographic images and T1- or T2-weighted images should be parameterized differently. To account for such differences, the authors propose a hierarchical arrangement of neural networks trained to provide accurate estimation for a wide range of images. To achieve the mapping of optimal transformation parameter values with individual images, a variety of histogram, wavelet and spatial features are used to characterize each image. The decision of each network module in the hierarchy is combined using a weighted averaged fusion scheme. Evaluating their framework on a dataset of more than 2400 images the authors report that their methodology gives satisfactory results and it is robust to unknown images.

Similarly, Pitiot *et al.* [13] proposed an automated method for extracting anatomical structures in MRI based on textural classification. The authors hypothesize that performing a pretextural classification prior to segmentation will

lead to a more accurate definition of the anatomical boundary. The pretextural classification is based on a mixture-of-experts paradigm, such that each expert is trained on a particular grouping of textural features extracted from a moving window within the image. A second-stage multiscale neural network is trained on equally drawn numbers of random samples from correctly and incorrectly classified pixels from the first stage. The network arrangement of stage two is trained on local morphology and texture features from a wider pixel neighborhood in the task of detecting anatomical structures. Evaluating their framework on a small dataset of 10 testing images, the authors report an increase in classification rates as a result of the two-stage hybrid neural classification.

Sha and Sutton [14] proposed the use of a network-of-networks paradigm first discussed by Guan *et al.* [17] for dynamically reconfiguring a test image for enhancement and segmentation. Under the proposed framework, the image is connected on a pixel-by-pixel basis by weights in a manner analogous to an attractor neural network. Pixels are updated based on local variances obtained from weight connections with neighbours in an iterative manner until convergence of the network architecture. In their study, the authors present only qualitative results on the enhancement and segmentation of MRI images.

11.2.1.3 Knowledge Representation Learnt from User Interactions

A novel method of implementing an adaptable characteristic in a knowledge-based component has been suggested by capturing user interactions with a CAD tool.

Fenster and Kender [15] proposed the use of a diagnostic tool for the interpretation of computed tomography (CT) images. The authors utilize a boundary-based segmentation technique termed the live wire paradigm. The motivation for the scheme is based on an attempt to utilize the interaction with an expert clinician during segmentation, thereby learning from the users feedback for use in subsequent segmentations. Under the proposed framework, a 2D boundary is constructed around a ROI based in part on image properties and knowledge acquired in training when manual segmentation is performed. The tool utilizes an optimal feature selection process to determine the best boundary position from the available information.

11.2.1.4 Knowledge Representation Using Case-Based Reasoning

Case-based reasoning (CBR) approaches have been used extensively as a means of directly utilizing image properties. CBR is an approach to computer-based cognition that involves reasoning from prior experiences. It solves new problems by adapting solutions that were used to solve old problems. The knowledge base of a CBR system consists of cases indexed by their pertinent features.

Perner [16] proposes a novel method of image segmentation based on CBR. The CBR unit for image segmentation consists of a case base in which formerly processed cases are stored. Accompanying each case is information regarding the parameters used in the segmentation of the image. On test, a similarity measure is used to select the most similar case in the archive from which the segmentation parameters are selected and used to segment the test image. The author hypothesizes that images having similar characteristics will show good segmentation results when the same segmentation parameters are applied. The image information comprises a vector of statistical features extracted from the gray scale histogram of the matching image. In the author's implementation, nonimage information is obtained from CT image headers, such as sex, age, CT-slice thickness, etc. The final similarity measure comprises both image and nonimage information. The segmentation of the image is performed using histogram analysis. The parameters for this process comprise those defining a smoothing function for the histogram and a set of thresholds used for histogram analysis. The system has been evaluated using a 130 image case base and by comparing the automatic CBR segmentation with that of an expert clinician on 600 CT images of the brain. The author reports a linear correlation coefficient of 0.85 between the CBR segmentation and those manually drawn by an expert clinician.

11.2.2 An Adaptive Knowledge-Based Model

As we can see from the survey in the previous section, little research has been undertaken in attempting to optimize a hierarchy of image processing operators. In this chapter an adaptive knowledge-based model is proposed. The model comprises deterministic and knowledge-based components for the detection of

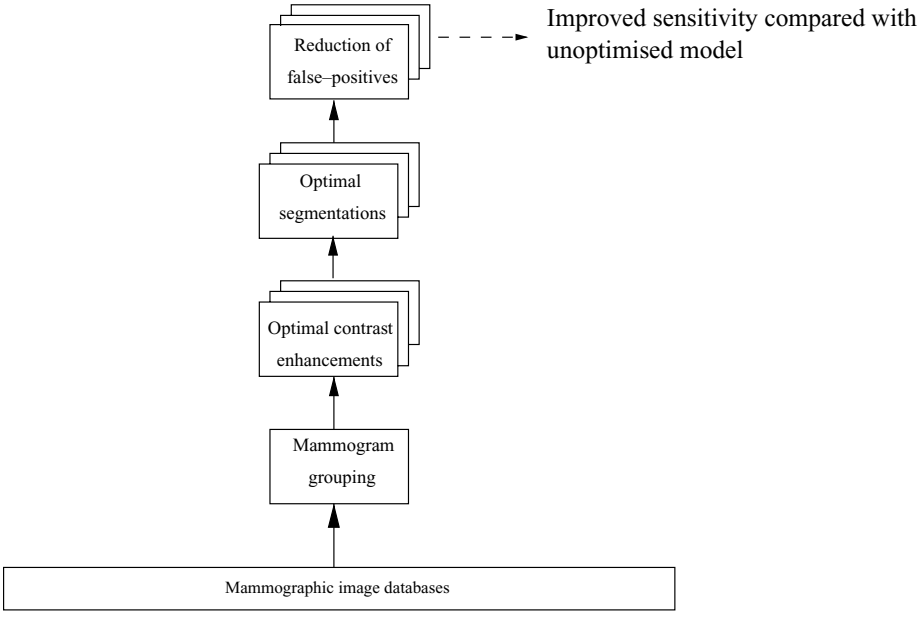


Figure 11.1: High-level flowchart identifying knowledge-based components to support the adaptive knowledge-based model.

breast cancer masses from screening digitized mammograms. An overview of the construction of the adaptive knowledge-based model is presented in this section.

11.2.2.1 High-Level Overview of Adaptive Knowledge-Based Model

Figure 11.1 shows a diagrammatic high-level overview of the proposed adaptive knowledge-based model. The major components of a CAD pyramid are shown. They are contrast enhancement, image segmentation for the identification of suspicious a ROI identification, and false-positive reduction. The knowledge-based framework underpinning the adaptive knowledge-based model is used in the identification of an optimal pipeline for each mammogram. Additional knowledge is incorporated into the model by implementing separate parameterized versions of image segmentation and false-positive reduction components according to a mammogram grouping strategy. Each knowledge-based component presented in Fig. 11.1 is discussed in further detail below.

Mammogram grouping: By grouping mammograms on a predefined criteria, subsequent CAD components may be engineered to operate on specific mammograms types. Our study hypothesizes that a mammogram can be grouped on the basis of its parenchymal patterns. The aim of this component is to predict a mammogram's group by utilizing supervised learning techniques in conjunction with a training set of example images.

Optimal contrast enhancement: A range of contrast enhancement techniques previously used in mammographic CAD research are surveyed in [18]. The adaptive knowledge-based model aims to accommodate many of these methods in the form of enhancement experts and learn, on the basis of feature vectors from training mammograms, the optimal contrast enhancement expert for a given mammogram on test. We propose machine learning techniques such as artificial neural networks (ANN) for learning this mapping.

Optimal image segmentation: A variety of different image segmentation methods have been identified for mammographic CAD in [18, 19]. Adopting a similar strategy to that of the knowledge-based contrast enhancement experts, a set of segmentation experts are proposed. As opposed to different contrast enhancement experts, each segmentation expert is functionally identical. The adaptability property in the segmentation component is achieved by learning the saliency of input features used to perform the segmentation. This chapter hypothesizes that different segmentation experts operating on different input feature spaces will have a greater utility in the segmentation of different mammograms. Input features for expert construction will be drawn from a subset commonly utilized in mammographic CAD. For example, the subset may include image gray scales, contrast enhanced gray scales, textures, and edge-gradient information possibly at different scales of resolution. Each segmentation expert operates on a predefined set of features for a predefined group of mammograms. The implementation of an optimal segmentation is achieved by predicting the best blend of segmentation decisions given by the collection of experts for an individual mammogram.

Reduction of false-positive regions: The final component is the reduction of false-positive regions. This component operates by discriminating between normal and abnormal regions based on a feature vector

extracted from a suspicious ROI. This component is implemented within the adaptive knowledge-based model as a modular arrangement of ANNs trained to specialize in particular groupings of mammograms.

11.2.2.2 The Complete Adaptive Knowledge-Based Model Framework

A schematic low-level representation of the adaptive knowledge-based model is shown in Fig. 11.2. Each component identified in Fig. 11.2 is discussed now to show how the complete model comprising a knowledge-based framework is to be implemented. Each blocked level in Fig. 11.2 is described in further detail in each part of this chapter and summarized below:

Image preprocessing component: The following image preprocessing is performed. Firstly, mammograms are grouped on the basis of their breast density. The Digital Database of Screening Mammograms (DDSM) database used in our study (http://marathon.csee.usf.edu/Mammography/Database.html) comes with the complete ground-truth definitions of breast cancer lesions specified according to the American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) lexicon. We develop a classification scheme that indexes the texture features of training data with their ground truth breast density information. This classification scheme is used to predict the correct density of a test image and hence used the correct algorithms for application.

Secondly, we redefine the boundaries (ground truth) supplied with the DDSM database to accurately represent the location of the lesion. This is done using an active contour model (full details are available in [18]). The reason for using this redefined ground truth is to improve the learning in the contrast enhancement and image segmentation components.

Expert image contrast enhancement: The main aim of this component is to select the optimal image enhancement method per image. The aim is to choose the enhancement method that optimizes the following image segmentation performance.

Expert image segmentation: The aim of image segmentation is to label pixels within an image as corresponding to real world objects. For a

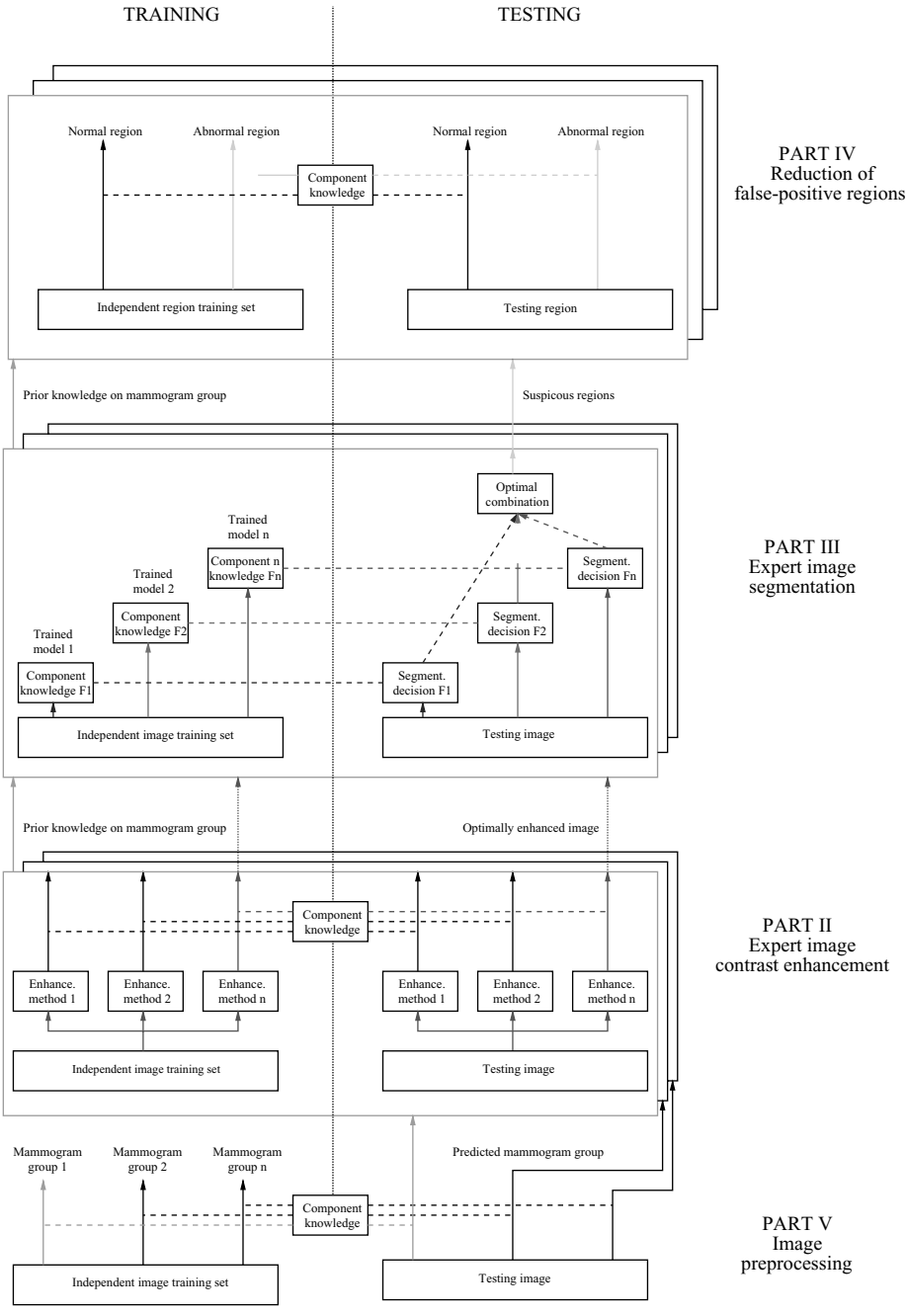


Figure 11.2: Process flowchart of proposed adaptive knowledge-based model. Principles of machine learning are used for training components on the left that are subsequently used for testing on the right. Flow of mammograms is from bottom up.

mammographic CAD scheme, this involves labeling pixels in the image as being normal or suspicious. In this way suspicious pixels may be combined into suspicious regions. By utilizing machine learning principles, a segmentation expert can be constructed from a set of training images drawn from a particular mammographic breast type. Our study evaluates the knowledge-based segmentation component using 10 different input feature spaces, including the original image, contrast-enhanced image, and a textural representation at different scales. To segment a mammogram, of a given breast type, the 10 trained segmentation experts each give an estimate of the segmentation based on their input feature spaces. These decisions are then combined such that the optimal blend of segmentation experts is determined thereby resulting in an optimal segmentation. From the segmented image, region boundaries are identified and the regions passed onto the final component for the reduction of false-positive regions.

Reduction of false-positive regions: We select a set of region-based features that can be used in conjunction with a separate training set of regions for learning component knowledge. By training an ANN for each breast type, a modular arrangement of ANNs can be used to specialize in decision-making. The aim on test is to reduce the average number of false-positive regions per image while maintaining a high level of sensitivity to lesion detection.

We now detail the individual components of the model in much greater detail.

11.3 Image Contrast Enhancement Layer

In order to construct a scheme for the optimal selection of image enhancement, some quantitative indices are needed that measure the amount of enhancement. Not enough research has been conducted to tackle this difficult issue. In our previous work [19, 20] we introduced three new quantitative measures of image enhancement based on the change in contrast between the target (mass) and the backgound (a border 20 pixels wide around the target). We cover these measures for the sake of completeness here in section 11.3.1. In addition, we also discuss

an independent measure of contrast called "difference in average separation" that has been popularly used in other work.

11.3.1 Measures of Contrast Enhancement

11.3.1.1 Distribution Separation Measure

Using the method for labeling the Target (T) and Background (B) regions, it is possible to plot the overlap of the density functions for the gray scales comprising these two regions. In mammography, this is representative of the overlap found between a breast cancer lesion and its background border. A good enhancement technique should ideally reduce the overlap. In particular, it is anticipated that the enhancement technique should help reduce the spread of the target distribution and shift its mean gray-scale level to a higher value thus separating the two distributions and reducing their overlap. The best decision boundary for the original image between the two classes, assuming both classes have a multivariate normal distribution with equal covariances, is given using [21] as

$$D_1 = \frac{\mu_B^{\mathrm{O}} \cdot \sigma_T^{\mathrm{O}} + \mu_T^{\mathrm{O}} \cdot \sigma_B^{\mathrm{O}}}{\sigma_T^{\mathrm{O}} \cdot \sigma_T^{\mathrm{O}}} \tag{11.1}$$

Similarly, the best decision boundary for the original image after enhancement is given as

$$D_2 = \frac{\mu_B^{\mathrm{E}} \cdot \sigma_T^{\mathrm{E}} + \mu_T^{\mathrm{E}} \cdot \sigma_B^{\mathrm{E}}}{\sigma_T^{\mathrm{E}} \cdot \sigma_B^{\mathrm{E}}} \tag{11.2}$$

where μ_B^{O}, σ_B^{O}, μ_T^{O}, and σ_T^{O} are the mean and standard deviation of the gray scales comprising the background and target area, respectively, of the original image before enhancement. Similarly μ_B^{E}, σ_B^{E}, μ_T^{E}, and σ_T^{E} correspond to the mean and standard deviation of the gray scales after the enhancement. An alternative approximation to D_1 and D_2 can be found using the cutting score [22]. If the groups are assumed to be representative of the population, a weighted average of the group centroids will provide an optimal cutting score where Eq. (11.1) is rewritten as

$$D_1 = \frac{\mu_B^{\mathrm{O}} \cdot N_T^{\mathrm{O}} + \mu_T^{\mathrm{O}} \cdot N_B^{\mathrm{O}}}{N_T^{\mathrm{O}} \cdot N_B^{\mathrm{O}}} \tag{11.3}$$

and Eq. (11.2) is rewritten as

$$D_2 = \frac{\mu_B^{\mathrm{E}} \cdot N_T^{\mathrm{E}} + \mu_T^{\mathrm{E}} \cdot N_B^{\mathrm{E}}}{N_T^{\mathrm{E}} \cdot N_B^{\mathrm{E}}} \tag{11.4}$$

where N_B^{O} and N_T^{O} are the number of samples in the background and target prior to enhancement, and N_B^{E} and N_T^{E} the respective sample numbers after the enhancement. Again this approximation assumes that the two distributions are normal and that the group dispersion structures are known. By combining the above two equations it is possible to compute a distance measure between the decision boundaries and the means of the targets and background, before and after segmentation. This measure is termed as the distribution separation measure (DSM), and it is a measure of the quality of enhancement. It is defined as

$$\mathrm{DSM} = \{|(D_2 - \mu_B^{\mathrm{E}}| + |(D_2 - \mu_T^{\mathrm{E}})|\} - \{|(D_1 - \mu_B^{\mathrm{O}}| + |(D_1 - \mu_T^{\mathrm{O}})|\} \tag{11.5}$$

Ideally the measurement should be greater than zero; the greater the DSM value, the better the quality of enhancement. For comparing any two enhancement techniques, choose the technique that gives a higher value on the DSM measure.

11.3.1.2 Target to Background Contrast Enhancement Measurement Based on Standard Deviation

A key objective of a contrast enhancement is to maximize the difference between background and target mean gray level and ensure that the homogeneity of the mass is increased aiding the visualization of its boundaries and location. Using the ratio of the standard deviation of the gray-scales within the target before and after the enhancement, the improvement using the target to background contrast enhancement using standard deviation ($\mathrm{TBC_s}$) is given as

$$\mathrm{TBC_s} = \left\{ \frac{(\mu_T^{\mathrm{E}}/\mu_B^{\mathrm{E}}) - (\mu_T^{\mathrm{O}}/\mu_B^{\mathrm{O}})}{\sigma_T^{\mathrm{E}}/\sigma_T^{\mathrm{O}}} \right\} \tag{11.6}$$

where the mean and standard deviation of the gray scales comprise the target and background before and after the enhancement. Assuming that the target has a smaller mean before and after enhancement compared to the background, it is expected that as a result of enhancement, this measure should give a value greater than zero.

11.3.1.3 Target to Background Contrast Enhancement Measurement Based on Entropy

It is possible to extend the concept of $\mathrm{TBC_s}$ further by replacing the standard deviation with the entropy of the target in the original and enhanced images, $\varepsilon_T^{\mathrm{O}}$ and $\varepsilon_T^{\mathrm{E}}$, respectively, to quantify the homogeneity ratio. Similar to Eq. (11.6) the target to background contrast enhancement using entropy (TBC_ε) is defined as

$$\mathrm{TBC_s} = \left\{ \frac{(\mu_T^{\mathrm{E}}/\mu_B^{\mathrm{E}}) - (\mu_T^{\mathrm{O}}/\mu_B^{\mathrm{O}})}{\varepsilon_T^{\mathrm{E}}/\varepsilon_T^{\mathrm{O}}} \right\} \tag{11.7}$$

Assuming that the target has a smaller mean before and after enhancement compared to the background, it is expected that as a result of enhancement, this measure should give a value greater than zero.

11.3.1.4 The Combined Enhancement Measure

It is possible to combine the three novel measures into a single quantitative value. Using this combined measure, a researcher is able to quantitatively rank enhancements for a particular image. To combine DSM, $\mathrm{TBC_s}$, and TBC_ε for a particular enhancement, each enhancement value is represented within a 3-D Euclidean space by min–max scaling each within the range [0,1]. A high performance contrast enhancement method will have points close to coordinates (1, 1, 1). The combined measure D is computed by calculating the Euclidean distance between the point in the 3-D coordinate space representing the enhancement and (1, 1, 1). This point in the enhancement measurement space represents the location of an enhancement method that results in the maximal increase in contrast between a target and its background. The combined measure D is computed as

$$D = \sqrt{(1 - \mathrm{DSM})^2 + (1 - \mathrm{TBC_s})^2 + (1 - \mathrm{TBC}_\varepsilon)^2} \tag{11.8}$$

The enhancement method giving the smallest value of D is selected as the best enhancement method for this image.

11.3.1.5 Difference in Average Separation Measure

This measure is defined as the difference in average separation (AVS) [23] between the original and corresponding enhanced image. The average

separation is a measure of intergroup dissimilarity and is defined as the average Euclidean distance d between "confused pixels," that is, pixels with the same gray scales found in both target and background regions. The AVS measure is defined as

$$\mathrm{AVS}(\omega_1, \omega_2) = \frac{1}{n_1 n_2} \sum_{i=1}^{n1} \sum_{j=1}^{n2} d(x_i, y_j) \qquad x_i \in \omega_1 \quad x_j \in \omega_2 \tag{11.9}$$

for all pairs of points such that a single point is drawn from each region, target ω_1 and background ω_2 with n_1 and n_2 pixels in total respectively. A large value of $\mathrm{AVS}_{\mathrm{diff}}$ will result if the enhanced image has a greater intergroup dissimilarity for gray scales in the target and background region compared with that of the original. This increased value of $\mathrm{AVS}_{\mathrm{enhanced}}$ indicates that the enhancement has maximized the Euclidean distance of the confused pixels thereby resulting in an improved contrast enhancement.

11.3.2 Contrast Enhancement Mixture of Experts Framework

In CAD of breast lesions, one aim of contrast enhancement is to improve the performance in image segmentation. Therefore, the optimal contrast enhancement method for a mammogram is the one maximizing the sensitivity of the detection of a breast lesion following image segmentation. The proposed knowledge-based enhancement component will predict the optimal contrast enhancement method, or expert, for a test mammogram using knowledge learnt from a set of training mammograms. Figure 11.3 summarizes the enhancement component framework. Individual enhancement experts $(1, \ldots, n)$ are used in training and testing, shown on the left and right of Fig. 11.3, respectively. Experts are grouped together in training and testing for particular mammogram types. During training, an optimal enhancement expert is identified (say for example expert 2) and the mapping between a global characteristic of the training mammogram and the enhancement expert is captured as component knowledge. For the testing mammogram, the optimal expert can be predicted (which was expert 2) based on an image feature vector. Where possible, the *a priori* knowledge of the mammogram breast type will be used. Different parameterized versions of the knowledge-based contrast enhancement component are constructed for each breast type grouping.

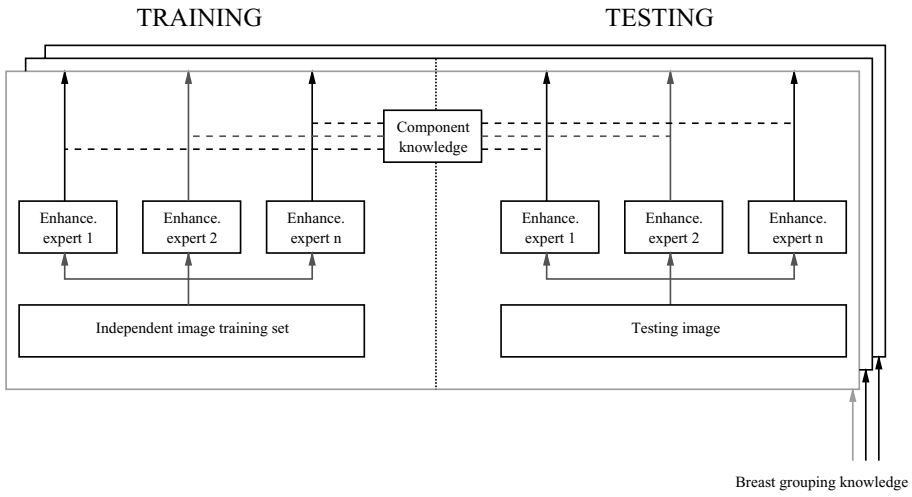


Figure 11.3:

This section describes the mixture of experts framework and it is laid out as follows. Section 11.3.2.1 reviews the contrast enhancement experts used to build the framework. Then the segmentation algorithm used to evaluate the enhanced images is briefly described together with quantitative measures of segmentation performance. In section 11.3.2.2 results are presented when applying the different image enhancement on DDSM images and the resulting segmentation from them. Section 11.3.2.3 discusses the features that can be extracted from the mammograms to be fed into a mapping scheme (e.g., neural networks) that maps features to optimal enhancement methods. Finally, section 11.3.2.4 discusses a machine learning system for this mapping. A neural network is used in two different modes: double network mapping and a single direct mapping scheme.

11.3.2.1 Segmentation of Contrast-Enhanced Digitized Mammograms

The aim of the knowledge-based contrast enhancement component is to predict the optimal contrast enhancement for a given mammogram. The optimal contrast enhancement is the one maximizing the segmentation of the enhanced image. Segmentation performance is measured by the sensitivity in the detection of true-positive regions within the segmentation image. Section 11.3.2.1.1 identifies

the contrast enhancement methods used in this configuration of the adaptive knowledge-based model. Following this, sections 11.3.2.1.2 and 11.3.2.1.3 describe the segmentation method used to evaluate the performance of the contrast-enhanced image and quantitative evaluation of segmentation quality. Finally, section 11.3.2.1.4 identifies the optimal contrast enhancement that exists for each mammogram, providing evidence for the construction of a knowledge-based enhancement component to predict the optimal enhancement method.

11.3.2.1.1 Contrast Enhancement Experts. The utility of six contrast enhancement methods are evaluated: histogram equalization (HISTOEQ), fuzzy enhancement (FUZZY), density weighted contrast enhancement (DWCE), adaptive contrast enhancement (ACE), adaptive contrast enhancement with local entropy (ACELE), and adaptive contrast enhancement with local fractal dimension (ACELFD). Each of these methods will be used as enhancement experts within this configuration of the knowledge-based framework. Full details on the their algorithms is available in [18].

11.3.2.1.2 Segmentation Methods. The aim of image segmentation is to label a pixel in an image as belonging to one of the known corresponding real world objects. In the detection of breast lesions in digitized mammograms, image segmentation results in contiguous areas or regions of pixels, labeled as normal or suspicious. For the purpose of evaluating image enhancement, we use an unsupervised Gaussian mixture model (GMM) and hidden Markov random field (HMRF_U) model of image segmentation proposed by Zhang *et al.* [24]. For ease of referencing, this shall be referred to as HMRF_U in the rest of this chapter. The HMRF_U segmentation method is used to segment contrast-enhanced images so that the performance of the contrast enhancement can be determined. The HMRF_U segmentation algorithm operates in an unsupervised manner. The only *a priori* knowledge required for the segmentation is the maximum number of classes, L, from which a pixel is labelled. By setting $L = 2$, HMRF_U will label pixels as either normal or suspicious. The HMRF_U method models each class using a single Gaussian whose parameters are defined using a maximum likelihood estimate. Followoing convergence, a maximum a posteriori (MAP) segmentation is performed by labeling each pixel with the class maximising the *a posteriori* probability estimates.

Table 11.2: Outcomes detected following image segmentation

Outcome	*Description*	*Summary*
TP	A detected area is defined as True-Positive (TP) if the following two conditions are true:	
	1. The common area between the Actual A and Target T divided by the area of the target region is greater than or equal to a certain percentage T_{min}.	1. $\frac{(A_{\text{area}} \cap T_{\text{area}})}{T_{\text{area}}} \geq T_{\text{min}}$
	2. The total area of the segmented actual region must be less than a constant C_{max}.	2. $(A_{\text{area}} \leq C_{\text{max}})$
	In this evaluation $T_{\text{min}} = 50\%$ and C_{max} is four times the size of the image target region in the complete dataset.	
SUBTP	A detected area is defined as SUBTP if the two following conditions are true.	
	1. The overlap area between the target and the actual regions is less than T_{min}.	3. $(A_{\text{area}} \cap T_{\text{area}}) < T_{\text{min}}$
	2. The actual area is less than or equal to C_{max}.	4. $(A_{\text{area}} \leq C_{\text{max}})$

11.3.2.1.3 Quantifying Segmentation Performance. To evaluate the quality of the segmentation of a mammogram enhanced using a particular contrast enhancement method, a mechanism for quantifying the segmentation performance is required. Utilizing a simple sensitivity outcome of detection will not identify the true segmentation performance. Instead, a measure of the area of the target ROI correctly labelled as suspicious is proposed. This is achieved using an overlap methodology described by Kallergi *et al.* [25]. In their study, the authors describe a series of quantitative measures based on the overlap of a suspicious actual region following image segmentation with that of a *target* ground truth region, denoted by an expert radiologist. Based on the study by Kallergi *et al.* [25], two quantitative measures are selected and described in further detail in Table 11.2. Figure 11.4 illustrates diagramatically the two outcomes

(a) (b)

Figure 11.4: Diagrammatic example of a (a) TP and (b) SUBTP.

to be detected. In each case the target region is shown as darker color and the actual region, following segmentation, is shown as lighter color overlapping. Figure 11.4(a) shows the TP outcome where the target and actual region overlap is greater than $T_{\min} = 0.5$ and conversely, Fig. 11.4(b) where the overlap of the target region is less than $T_{\min} = 0.5$, the SUBTP outcome.

11.3.2.2 Evaluation on DDSM Mammograms

This section presents the results obtained from the segmentation of 200 mammograms from the DDSM. The aim of the experiment is to identify the optimal contrast enhancement expert for each of the 200 abnormal mammograms. Each mammogram image has been grouped according to its target breast type. There are 50 images per breast type grouping and results will be presented on a per breast type basis. Each mammogram is contrast enhanced using each enhancement method identified in section 11.3.2.1.1 and segmented using the unsupervised HMRF_{U} segmentation method. The sensitivity in the detection of breast lesions following segmentation of the enhanced images is quantified using the outcomes given Table 11.2 and the ground truth definition.

From a set of M enhancement methods $(E_1, \ldots, E_M)$ for a given mammogram, the target contrast enhancement, E_m where $m \in \{1, \ldots, M\}$ is the enhancement method giving the largest value of $(\text{TP}^T + \text{SUBTP}^T)$ following segmentation using HMRF_{U}. The target contrast enhancement expert E_m is identified as

$$\text{assign } E_m \rightarrow m \text{ if } (\text{TP}^T + \text{SUBTP}^T)_m = \arg\max_{m=1}^{\text{M}} (\text{TP}^T + \text{SUBTP}^T)_m \tag{11.10}$$

Table 11.3: Segmentation results from using the original mammogram

Type	TP^T	$SUBTP^T$	Total
1	0.25	0.04	0.29
2	0.18	0.11	0.29
3	0.20	0.04	0.23
4	0.15	0.02	0.16
Mean	0.20	0.05	0.24

Values given are mean percentage of mass detected with TP and SUBTP outcome together with their sum.

The target contrast enhancement E_m is found for every mammogram from all M enhancement methods, $m \in \{1, \ldots, M\}$, keeping the segmentation method and associated initialization parameters constant. Having identified each of the target enhancement experts, the following important observations can be made (see sections 11.3.3.2.1–11.3.2.2.6).

11.3.2.2.1 Original vs. Contrast-Enhanced Mammograms. Each original unenhanced image is segmented and out of the 200 mammograms used, 80 (40%) images give no sensitivity in the detection of target regions, that is, for these images the value of $(TP^T + SUBTP^T) \leq 0$. Of these 80 unenhanced images, 34 still give no sensitivity in detection after application of each evaluated contrast enhancement method. Only the remaining 166 abnormal mammograms are considered in the evaluation of the optimal strategies described in the following sections. Table 11.3 presents the results from segmenting the original unenhanced mammograms grouped by breast type. The table shows the mean percentage of target region detected with each outcome, grouped by breast type. We observe that the segmentation performance decreases as the breast density increases.

11.3.2.2.2 Improved Segmentation in Contrast-Enhanced Mammograms. Of the 200 mammograms used in this evaluation, 150 (75%) give a greater sensitivity after the target, contrast enhancement method, compared with the sensitivity obtained from the original images.

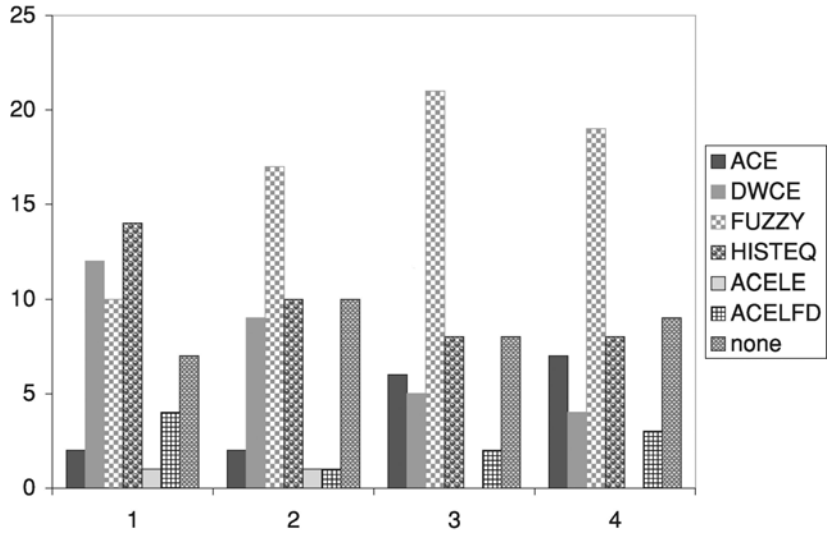


Figure 11.5: Frequency of each contrast enhancement method being selected as the target method for each mammogram grouped by breast type.

11.3.2.2.3 Reduced Sensitivity in Enhanced Images. Of the 166 enhanced mammograms that give a positive sensitivity result, 12 (7%) reported an inferior (reduced) sensitivity on application of the target contrast enhancement method compared with the original mammogram. No attempt is made to learn that "no contrast enhancement" is best suited for these images because of the small sample number.

11.3.2.2.4 Frequency of Optimal Contrast Enhancement Methods. The target contrast enhancement for a given mammogram is defined above using Eq. (11.10). Figure 11.5 presents the frequency that each of the six enhancement method are identified as the target contrast enhancement for a mammogram, grouped according to its breast type. From this figure, it can be seen that each enhancement methods is a target optimum for at least one mammogram. The variability in choice of target method is greater in the fatty breast (breast types 1 and 2) cases compared to the dense types (types 3 and 4). Note that the proposed novel extension to the ACE combining fractal dimension, method ACELFD, outperforms the classic ACE method more frequently for the fatty breast types. Additionally, it should be noted that the HISTOEQ method outperforms all other

Table 11.4: Mean percentage improvement in segmentation performance by using the target-enhancement method compared to the segmentation of the unenhanced image for each mammogram grouped by breast type

Type	TP^T	$SUBTP^T$	*Total*
1	0.76	1.25	0.83
2	1.28	0.64	1.03
3	0.80	4.00	1.43
4	1.47	6.50	2.25
Mean	1.07	3.09	1.38

methods for fatty breasts but is noticeably less effective in the segmentation of dense breasts (types 3 and 4).

11.3.2.2.5 Mean Percentage of Target Mass Detected as TP^T or $SUBTP^T$. Using the target contrast enhancement expert, Table 11.4 tabulates the mean percentage improvement in segmentation performance compared with the segmentation of the unenhanced image for lesion ground-truth detected with outcomes TP^T and $SUBTP^T$ for all mammograms, grouped by breast type. The greatest improvement can be seen in the segmentation of the densest breast type using the target contrast enhancement method, compared to the segmentation obtained from the unenhanced original (e.g., on average, for breast type 1, using target enhancement method result in 83% improvement in segmentation compared to the unenhanced original).

11.3.2.2.6 A Classical Solution in Choosing a Single Optimal Contrast Enhancement Method for CAD. In analyzing the 166 selected mammograms, a common approach to identify a single optimal contrast enhancement for use in CAD is to evaluate a selection of enhancement methods on a range of different training mammograms from different breast types. By determining the mean value of $TP^T + \text{SUBTP}^T$ for all enhancement methods across all mammograms, the CAD researcher can choose to select the single method maximizing the value of $\text{TP}^T + \text{SUBTP}^T$ from the training set. Table 11.5 lists the

Table 11.5: Mean percentage improvement in segmentation performance for each contrast enhancement method compared with the segmentation of the enhanced image, grouped by breast type

Type	TP^T	$SUBTP^T$	Total
(a) ACE			
1	−0.24	0.00	−0.21
2	0.17	−0.55	−0.10
3	0.10	−0.25	0.09
4	−0.07	0.00	0.00
Mean	**−0.01**	**−0.20**	**−0.06**
(b) DWCE			
1	0.12	0.00	0.10
2	−0.22	−0.27	−0.21
3	0.05	0.25	0.13
4	−0.13	1.00	0.06
Mean	**−0.05**	**0.24**	**0.02**
(c) FUZZY			
1	−0.12	2.50	0.24
2	0.17	1.00	0.48
3	0.05	5.50	1.04
4	0.33	8.00	1.38
Mean	**0.11**	**4.25**	**0.79**
(d) HISTEQ			
1	−0.28	0.00	−0.24
2	0.50	−0.55	0.07
3	0.00	−0.50	−0.04
4	0.20	1.00	0.38
Mean	**0.11**	**−0.01**	**0.04**
(e) ACELE			
1	−0.24	0.75	−0.10
2	0.06	−0.45	−0.17
3	0.10	−0.25	0.09
4	−0.20	0.50	−0.06
Mean	**−0.07**	**0.14**	**−0.06**
(f) ACELFD			
1	−0.24	0.25	−0.17
2	−0.11	−0.55	−0.28
3	0.05	0.25	0.09
4	−0.07	0.50	0.06
Mean	**−0.09**	**0.11**	**−0.07**

percentage improvement in segmentation compared with the original segmentation, in the detection of ground truth defined lesions detected with outcomes TP^T and $SUBTP^T$ obtained using each individual enhancement method. From the data obtained, in identifying the single best enhancement method, the FUZZY method (Table 11.5, part c) is chosen for each breast type. Notice that the segmentation performance obtained using the FUZZY method shown in Table 11.5 (part c) for all mammograms, is suboptimal compared with that of the target contrast enhancement for each mammogram shown in Table 11.4.

11.3.3 Identifying Input Mapping Features

To implement a knowledge-based contrast enhancement component to learn the target enhancement for a given mammogram, a supervised learning paradigm is employed. By utilizing pattern recognition tools, a classifier can learn from a set of example mammograms the target contrast enhancement. For an unseen testing mammogram, the trained classifier will accurately predict the actual enhancement that maximizes segmentation performance.

During training the classifier learns a mapping between a characteristic of an example training mammogram and the target enhancement method. To facilitate this mapping, features are extracted to characterize the training and testing mammograms. Two different approaches to feature extraction are described: (1) feature extraction from a ROI and (2) feature extraction from a breast profile.

11.3.3.1 Gray-Scale Features Extracted from a Suspicious ROI

This approach to feature extraction extracts a set of F features from pixels comprising a suspicious ROI target, T, thus $F_{ROI} = \{f_1, f_2, \ldots, f_F\}$. A surrounding region labelled background, B, of the same area is constructed encircling the ROI, but comprising normal pixels. From pixels that comprise the target (T) and background (B) region, the following gray-level statistics are extracted: *mean*, *standard deviation*, *entropy*, *skewness*, and *kurtosis*. These are transformed into feature values by determining the ratio of the target value to background value (T/B) for each gray-scale statistic. The features reflect the mathematical composition of the quantitative measure of contrast enhancement previously proposed.

11.3.3.2 Gray-Scale Features Extracted from the Breast Profile

An alternative approach to feature extraction is based on the gray scales that comprise a breast profile. A method for the construction of a segmented breast profile is described in [19]. A number of gray-scale features are extracted from the mammograms, including co-occurrence matrix based features, Fourier Transform-spectral energy based features, Law's mask features, discrete wavelet transform features, statistical features, circularity (shape) feature and fractal dimension feature. These are described in detail in [26]. Altogether this gives a set of 316 features $F_{\mathrm{BP}^{316}} = (f_1, f_2, \ldots, f_{316})$ that are extracted from the breast profile to characterize a mammogram. In addition, the application of principal component analysis (PCA) (a mechanism for dimensionality reduction) results in a 26-dimensional feature space $F_{\mathrm{BP}^{26}} = (f_1, f_2, \ldots, f_{26})$. Both features sets $\{F_{\mathrm{BP}^{316}}, F_{\mathrm{BP}^{26}})$ are evaluated as mapping features in the learning of the expert contrast enhancement.

11.3.4 Strategies for Learning the Contrast Enhancement Experts

To train a knowledge-based contrast enhancement component, the mapping between the input gray-scale feature vectors discussed in the previous section and the target method indicated by the quantitative measure of segmentation from Table 11.2 is learnt. This section gives an overview of two strategies to learn the target enhancement mapping:

Double network mapping (DNM): This strategy adopts a divide-and-conqueror paradigm. It attempts to decompose a single mapping into two simpler mappings. The first mapping to be learnt between the features from ROI and the three quantitative measure of enhancement performance proposed in section 11.3.1. A second process learns the mapping of the quantitative measure of enhancement with quantitative measure of segmentation. On testing this strategy will predict a measure of segmentation for each contrast enhancement method, and the actual contrast enhancement method is identified as the one maximizing the segmentation performance.

Breast profile mapping (BPM): This strategy differs in that the solution adopts a classification-based approach and aims to learn the mapping of feature set F_{BP} extracted from the complete breast image with a label of the target

enhancement. On test, a single contrast enhancement method is predicted. Each strategy is described in detail in the following sections.

11.3.4.1 Double Network Mapping Overview

The double network mapping (DNM) method is used to predict the target contrast enhancement using two ANNs for each enhancement method. The aim is to learn a mapping based on a set of gray-scale features F_{ROI} from a given mammogram, with a quantitative measure of segmentation performance, S. The segmentation performance is quantified following contrast enhancement, for each enhancement method m, where $1 \leq m \leq M$ from a set of M enhancement methods. The two submappings are detailed below:

1. ANN^{m}_{DNMenh}: For a mammogram I, enhanced using enhancement method m, this ANN learns the mapping between the set of F gray-scale input features $F_{\mathrm{ROI}} = (f_1, f_2, \ldots, f_F)$ extracted from a suspicious ROI, and a set of P quantitative measures $Q = (q_1, q_2, \ldots, q_P)$ of enhancement performance as described previously in section 11.3.1.

2. ANN^{m}_{DNMseg}: For a mammogram I, enhanced using enhancement method m, the ANN learns the mapping between the set of quantitative measure $Q = (q_1, q_2, \ldots, q_P)$ of enhancement performance and a set of R measures quantifying the performance of lesion segmentation $S = (S_1, S_2, \ldots, S_P)$ identified in Table 11.2.

A diagrammatic overview of the mappings learnt is given in Fig. 11.6 and the training and testing phases are described in more detail below. To evaluate the strategy, a firefold cross-validation approach is used to reduce bias and ensure that a test result is produced for each mammogram image.

11.3.4.1.1 Training the DNM Approach. Using this strategy, ANN^{m}_{DNMenh} and ANN^{m}_{DNMseg} are trained independently for each enhancement method, E_m where $m \in \{1, \ldots, M\}$. For a training image, a border comprising normal pixels of the same areas as the target ROI is constructed around it. The set of gray-scale input features F_{ROI} are extracted from the target ROI and background regions as described in section 11.3.3.1. Each training mammogram is contrast enhanced with each method and a set of quantitative measures of enhancement

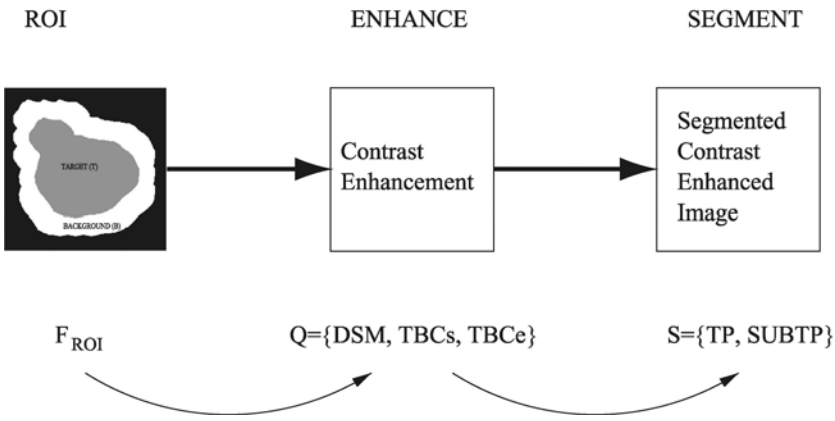


Figure 11.6: Diagrammatic overview of the DNM strategy.

Q are calculated from the target ROI and border. Thus ANN^m_{DNMenh} learns the mappings:

$$F_{ROI} \overset{ANN^m_{DNMenh}}{\longrightarrow} Q \quad \forall m = \{1, \ldots, M\} \tag{11.11}$$

Similarly for enhancement method E_m to train ANN^m_{DNMseg}, the set of quantitative measures of enhancement Q are used as input features in learning the mapping with the set of quantitative measures of segmentation, S. Thus ANN^m_{DNMseg} learns the mappings:

$$Q \overset{ANN^m_{DNMseg}}{\longrightarrow} S \quad \forall m = \{1, \ldots, M\} \tag{11.12}$$

11.3.4.1.2 Testing the DNM Approach. The first step in determining the optimal contrast enhancement method for a test mammogram I is to locate a suspicious ROI. To do this, the $HMRF_U$ segmentation algorithm is used to segment the test image and it results in several candidate regions. Regions with a *Euler* number > 1 (i.e, enclose a smaller region totally) are removed and from the remaining regions, the most likely suspicious regions are selected on the basis of area and morphological tests using a previously trained ANN. For the single suspicious ROI identified, a surrounding border is constructed of equal area, and the set of input gray-scale features F_{ROI} are extracted. These are used as inputs to ANN^m_{DNMenh} for each enhancement method E_m where $m \in \{1, \ldots, M\}$. The output of these networks is then supplied as input to ANN^m_{DNMseg} for each enhancement

Table 11.6: Results from using optimized strategy DNM showing the mean percentage improvement in segmentation performance compared with the segmentation of the unenhanced original mammogram

Type	TP_{μ}^{A}	$SUBTP_{\mu}^{A}$	Total
1	−0.16	1.75	0.10
2	0.28	−0.09	0.14
3	0.20	1.00	0.39
4	0.00	2.00	0.31
Mean	0.08	1.16	0.24

method $m \in \{1, \ldots, M\}$. The mth ANN_{DNMseg}^{m} maximising the value for the sum of TP and $SUBTP$ outcomes is predicted as the optimal enhancement method for a test mammogram I, thus

$$\text{assign } E_m \rightarrow m \text{ if } ANN_{DNMseg}^{m} = \text{argmax}_{m=1}^{M} ANN_{DNMseg}^{m} \text{ for } \forall m = \{1, \ldots, M\} \tag{11.13}$$

11.3.4.1.3 Model Order Selection. Optimization of the model order for ANN_{DNMseg}^{m} and ANN_{DNMseg}^{m} for each enhancement method E_m, where $m \in \{1, \ldots, M\}$, is performed independently by varying the number of hidden nodes between 2 and 30. The mean squared error (MSE) resulting for each model on test using fivefold cross validation is minimized. The optimal number of hidden nodes selected is the one which minimizes the MSE over all configurations of hidden nodes.

11.3.4.1.4 DNM Framework Results. Table 11.6 shows the mean percentage improvement in segmentation performance compared with the segmentation of the unenhanced original image, using the predicted expert enhancement for each breast type. The DNM strategy results are significantly poorer than those obtained using the target expert contrast enhancement methods reported in Table 11.4. They are also inferior to the use of the FUZZY method on all breasts as shown in Table 11.5 (part c).

11.3.4.2 Breast Profile Mapping—Overview

The second strategy used for learning the expert contrast enhancement for a mammogram is the breast profile mapping (BPM) strategy. For a mammogram I, enhanced using enhancement method E_m, where $m \in \{1, \ldots, M\}$, the BPM strategy learns the mapping between the set of N gray-scale input features F_{BP^N} detailed in section 11.3.3.1, and a $l = \{1., \ldots, L\}$ indicates the target contrast enhancement for a training mammogram. Both feature sets $\{F_{\text{BP}^{316}}, \{F_{\text{BP}^{26}}\}$ are evaluated separately in their utility for learning the expert contrast enhancement. The expert l is based on a set of R measures quantifying the performance of lesion segmentation $S = \{s_1, s_2, \ldots, s_R\}$ described in Table 11.2. The expert l is identified as the one maximising the sum of TP^T and SUBTP^T outcomes for each enhancement method E_m where $m \in \{1, \ldots, M\}$ as defined previously in Eq. (11.10).

Unlike the DNM strategy, this method utilizes a single classifier to predict the target contrast enhancement method. The k-nearest neighbor (k-NN) classifier has been show to be effective at learning nonparametric mappings with a small sample size [27] and for this reason it is employed in the knowledge-based contrast enhancement expert. To evaluate the strategy a five-fold cross validation is used to reduce bias and provide a test decision for each mammogram.

11.3.4.2.1 Training the BPM Approach. To train the BPM strategy, the set of gray-scale input features $F_{\text{BP}^N} = (f_1, f_2, \ldots, f_N)$, where N identifies the original and PCA feature sets ($N = \{316, 26\}$), are extracted from the segmented breast profile. Each training mammogram is contrast enhanced with each enhancement method. The quantitative measures of segmentation are calculated for the target ROI. For each enhancement method, the winning predicted enhancement method identified by the label l is used to learn the mapping between F and l with the k-NN classifier.

11.3.4.2.2 Testing the BPM Approach. To determine the predicted target enhancement method E_l for a test mammogram I, the set of gray-scale input features F_{BP^N} are extracted from the segmented breast profile. Using the trained k-NN classifier, the predicted actual expert contrast enhancement is determined.

Table 11.7: Percentage improvement in segmenting an unenhanced mammogram compared to that obtained when segmenting the image enhanced using the predicted enhancement method from the optimized BPM strategy based on feature set $F_{\mathrm{BP}^{316}}$

Type	TP_{μ}^{A}	SUBTP_{μ}^{A}	Total
1	−0.16	1.50	0.07
2	0.11	0.91	0.41
3	0.10	4.75	0.96
4	0.47	7.50	1.44
Mean	0.13	3.66	0.72

11.3.4.2.3 Model Order Selection. In order that the BPM strategy is to perform optimally, the number of nearest neighbors k must be correctly set. For each input feature set $F_{\mathrm{BP}^{316}}$ and $F_{\mathrm{BP}^{26}}$ for different values of k the validation set error is plotted and the value of k corresponding to the least error is chosen.

11.3.4.2.4 BPM Framework Results

1. Feature set $F_{\mathrm{BP}^{316}}$: Using an optimized value of k = 23, Table 11.7 shows the percentage improvement in segmentation performance when using the predicated actual enhancement method, compared with that obtained with the unenhanced original from the $F_{\mathrm{BP}^{316}}$ set. These results show that the segmentation improvement obtained over the unenhanced image, when segmenting an image enhanced using a enhancement method predicted by the BPM strategy, is greater than that obtained using the DNM strategy predicted enhancement method. However, segmenting the BPM strategy's predicted enhanced image results in inferior performance to that using the target enhancement method identified in Table 11.4. The result for breast type 4, the densest breast type, shows a small improvement over using the FUZZY method, shown in Table 5 (part c), for all mammograms of that type.

2. Feature set $F_{\mathrm{BP}^{26}}$: Using an optimized value of $k = 19$, Table 11.8 shows the percentage improvement in segmenting the unenhanced image compared

Table 11.8: Percentage improvement in segmenting an unenhanced mammogram compared to that obtained when segmenting the image enhanced using the predicted enhancement method from the optimized BPM strategy based on feature set $F_{\mathrm{BP}^{26}}$

Type	TP_{μ}^{A}	SUBTP_{μ}^{A}	Total
1	0.04	2.00	0.31
2	0.33	1.00	0.59
3	0.30	5.00	1.17
4	0.13	7.00	1.06
Mean	0.20	3.75	0.78

to that when segmenting the image enhanced using the predicted enhancement method by the optimized BPM strategy with the $F_{\mathrm{BP}^{26}}$ feature set. These results indicate better performance than the DNM strategy but are still inferior to the segmentation using the target expert enhancement method shown in Table 11.4. The result for breast type 1–3 show an improvement over using the FUZZY method in Table 5 (part c) for all mammograms of that type. Comparing the results from the evaluation of the two feature sets, $F_{\mathrm{BP}^{26}}$ and $F_{\mathrm{BP}^{316}}$ from the BPM strategy, the results indicate that the feature set $F_{\mathrm{BP}^{26}}$ is better suited to processing mammograms with breast types 1–3, whereas the feature set $F_{\mathrm{BP}^{316}}$ gives better performance on the densest breast type, i.e., type 4. Interestingly, for both feature sets, the performance improvement is worse over the fattiest breast types, type 1, compared with the densest, type 4. This is because of the variability of optimal enhancement method for the fatty breast types, whereas the denser breasts tend to be optimal enhanced by the FUZZY method more often.

11.3.5 Key Observations

Table 11.9 shows the shows the percentage improvement in segmenting the unenhanced image compared to that when segmenting the image enhanced using the predicted enhancement method for each of the two strategies compared that

Table 11.9: Mean percentage improvement in segmenting an unenhanced mammogram compared to that obtained when segmenting the image enhanced using the predicted enhancement method from each strategy for all breast types

Type	Mean TP	Mean SUBTP	Total
Target expert	1.00	2.00	1.20
FUZZY expert	0.11	4.25	0.79
DNM	0.08	1.16	0.24
(A) BPM $F_{\text{BP}^{26}}$	0.20	3.75	0.78
(B) BPM $F_{\text{BP}^{316}}$	0.13	3.66	0.72
Types 1–3 (A); Type 4 (B)	0.29	3.88	0.88

of the target optimal values from Table 11.4. Additionally, the table shows the result obtained by applying the FUZZY method to all images (given in Table 5(part c) over all four breast types. The last row in Table 11.9 shows the result of using the prediction from the BPM strategy with feature set $F_{\text{BP}^{26}}$ on breast types 1–3 and feature set $F_{\text{BP}^{316}}$ on type 4. From these results the following key observations are made:

1. *Utility of contrast enhancement*: From the complete dataset of mammograms, 75% showed an improved sensitivity following application of the expert contrast enhancement compared with the unenhanced original images.

2. *Target experts*: Figure 11.5 highlighted that given a set of contrast enhancement methods, different methods can be identified as target enhancement experts for different mammograms. This observation is the motivation for learning the optimal expert.

3. *Characterizing a mammogram*: Reviewing the results in Table 11.9, it can be seen that as the DNM strategy relies on characterizing a mammogram by a suspicious *ROI*, it performs poorly. In contrast the BPM strategy utilizes an image feature vector extracted from the breast comprizing an extensive set of features and performs better.

4. *The superior BPM approach*: The resultant performance using the modified BPM strategy based on breast type leads to a greater performance

than simply using the FUZZY method. The result is inferior to the target contrast enhancement baseline performance indicating that learning the expert enhancement is a nontrivial problem. In implementing the modified BPM strategy, a mechanism of predicting the breast type is required.

5. *Use of mammogram grouping knowledge*: The BPM approach has been developed to utilize *a priori* knowledge describing the mammogram grouping indicating the mammographic breast density type. This knowledge is used to determine the feature extraction method to be used, either $F_{\mathrm{BP}^{26}}$ for breast types 1–3 or $F_{\mathrm{BP}^{316}}$ for type 4. In the experimental results presented above, the target breast type was used.

11.4 Image Segmentation Layer

The image segmentation layer aims to use a number of image segmentation schemes and then adopt a mixture of experts model. In other words, on a per pixel basis, a number of segmentation experts make classification decisions that are fused together. The fusion of decisions is possible either using standard combination rules or adaptable scheme (based on determining appropriate weights of combination that are based on image properties). Our approach is based on the use of parametric models of image segmentation.

Recently, GMM have gained considerable prominence in the image segmentation literature since there is a vast range of training data available from which *a priori* information can be gathered. One of their key strengths is that such statistical models are underpinned by well-founded statistical probability and information theory. In addition, such approaches can be used in supervised or unsupervised modes. In addition, the output of such models is the *a posteriori* probability estimate that can be used to optimize the model to perform at a given point on the ROC curve. Also, by expressing the result as *a posteriori* probability, the outputs of various experts can be combined within a unified framework. Finally, the postprocessing of images is cheaper with statistical methods since only those regions that contain suspicious pixels need further examination, as opposed to a region-based approach where all regions must be considered.

The GMM approach does not consider the spatial arrangement of class labels in an image, which can be quite useful for relaxation labeling [28]. Markov random fields (MRF) have been shown as a powerful class of techniques [29–31] for modeling the spatial arrangement of class labels. MRF can be expressed in terms of a probabilistic framework and they can be combined with a statistical observed model of the mammogram. An MRF can increase the homogeneity of the formed regions that leads to a reduction in the false positives.

In this study we propose a Weighted Gaussian Mixture Model (WGMM) for both supervised ($WGMM_S$) and unsupervised ($WGMM_U$) data analysis. A set of GMMs is constructed, each modeling a particular class distribution and capable of being combined into a single unconditional density. We combine the WGMM model with a MRF hidden model and propose two approaches that work for supervised ($WGMM_S^{MRF}$) and unsupervised ($WGMM_U^{MRF}$) modes. The four models or experts ($WGMM_S$, $WGMM_U$, $WGMM_S^{MRF}$, and $WGMM_U^{MRF}$) each produce a label for the test pixel. We use a number of different features, each forming the basis of a different expert and relying on one of the above four models for segmentation. The expert outputs can be combined using well-known expert combination methods. In this chapter we propose an adaptive weighted model (AWM) for the combination of four experts and show that this new method of combination outperforms other popular methods.

11.4.1 Weighted Gaussian Mixture Models

A gray-scale image is represented as a 1-D array $X = \{x_1, x_2, \ldots, x_N\}$, where x_n is an input feature for pixel n and N is the total number of pixels in the image. The input feature vector x_n may be a D-dimensional vector or simply the gray-scale value of the pixel n. Let the underlying true segmentation of the image be denoted as $Y = \{y_1, y_2, \ldots, y_N\}$. It is assumed that the number of classes is predetermined as a set of known class labels ω_l, where $l \in \{1, \ldots, L\}$, and therefore the class label of pixel n is indicated as $y_n \in \{\omega_l\}_{l=1}^{L}$. A common assumption in modeling a density with a GMM for image segmentation is that each component m, $m \in \{1, \ldots, M\}$, will model the *pdf* of each class $M = L$. Let $\hat{y}_n$ represent the estimate of the segmentation. Each component is weighted by its weight of Y_{mn} that indicates the relationship of pixel x_n to class label ω_l modeled by component m. To ensure that the parameters of each component density are learnt correctly, the weight Y_{mn} is set to indicate the class to which

data point x_n belongs, thus

$$\gamma_{mn} = \begin{cases} 1 & \text{if } y_n = m \\ 0 & \text{otherwise} \end{cases}$$

If $Y_{mn} = 1$, then data point x_n will only be considered when setting the parameters of class ω_l modeled by component m. Using the labelled training data, a maximum likelihood (ML) estimate of all component parameters and mixing coefficients can be found.

We first describe the two modes of test image segmentation, supervised and unsupervised, in section 11.4.2. We then detail our weighted GMM/MRF models in section 11.4.3.

11.4.2 Supervised and Unsupervised Test Image Segmentation

A test image to be segmented is represented in the same way as the training image by a 1-D array X. In the case of test image, a 1-D array $\hat{Y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N\}$ is the estimate of the segmentation. We can now adopt one of the two strategies for test image segmentation.

1. *Supervised segmentation with GMM*: Using the ML estimate of the parameter values obtained from the training images, a segmentation of the test images is performed. This is achieved by substituting the learnt model parameters θ from training when performing testing. The image is segmented by setting the class label estimate $\hat{y}_n$ of pixel x_n as the one with the maximum estimate of the component-conditional probability.

$$\hat{y}_n = \arg \max_{m=1}^{M} \{p(y_n = m \mid x_n, \theta_m)$$

2. *Unsupervised segmentation with GMM*: This alternative approach assumes no *a priori* knowledge except for the number of classes in the image corresponding to the number of components in the GMM, $L = M$. Therefore, the weight $Y_{mn} = 1$ indicates that all samples are considered as being generated from this distribution. Using the GMM-EM algorithm, an ML estimate of the parameter values is found. The segmentation can then be estimated using the GMM by extracting the component-conditional probabilities using the Bayes rule.

11.4.3 A Weighted GMM/MRF Model of Segmentation

A finite mixture model (FMM) [23, 27, 32] is defined as a linear combination of M component conditional densities $f(x \mid m, \theta_m)$, for $m = \{1, \ldots, M\}$, and M mixing coefficients $f(m)$ of the form

$$f(x) = \sum_{m=1}^{M} f(m) f(x \mid m, \theta_m) \tag{11.14}$$

such that the mixing coefficients $f(m)$ satisfy the following constraints:

$$\sum_{m=1}^{M} f(m) = 1 \quad \text{and} \quad 0 \leq f(m) \leq 1.$$

The framework of WGMM comprises of $l \in (1, \ldots, L)$ class densities each modeled independently using a GMM of the form given in Eq. (11.14) and a set of mixing coefficients $p(\omega_l)$ as

$$p(x) = \sum_{l=1}^{L} p(\omega_l) p(x \mid \omega_l, \Theta_l) \tag{11.15}$$

The lth GMM estimates the class-conditional *pdf* $p(x \mid \omega_l, \Theta_l)$, which is itself another mixture model, for each data point for each class $\{\omega_l\}_{l=1}^{L}$. The vector Θ_l is defined as the M component Gaussian parameters of the lth *GMM* as $\Theta_l = \{P_l(m), \mu_{lm}, \Sigma_{lm}\}$, $\forall m = \{1, \ldots, M\}$. Each estimate of the class conditional *pdf* is mixed to model the overall unconditional density $p(x)$, using a mixing coefficient $p(\omega_l)$, identifying the contribution of the lth class density in the unconditional *pdf*.

If it is assumed that for a complete dataset X, of points x_n, where $X \equiv \{x_1, \ldots, x_N\}$, is drawn independently from the distribution $f(x \mid \theta)$, then the joint occurrence of the whole dataset can be conveniently expressed as the log likelihood as follows:

$$\log \zeta(\Theta) = \sum_{n=1}^{N} \log p(x_n \mid \Theta) = \sum_{n=1}^{N} \log \sum_{l=1}^{L} \gamma_{nl} p(\omega_l) p(x_n \mid \omega_l, \Theta l) \tag{11.16}$$

Using a modified version of the expectation-maximisation (EM) algorithm, as described below, we derive an ML estimate of the parameter values of each of the L GMMs $\{\Theta_l\}_{l=1}^{L}$.

The general framework for parameter estimation in GMM can be used to learn the parameters of WGMM. Here the component conditional densities, appearing

in Eq. (11.13) are themselves mixture models. In the EM algorithm, the update equations for mixing coefficients do not depend on the functional particulars of the component densities. Hence, the mixing coefficients of the WGMM are updated according to

$$P^{\text{new}}(\omega_l) = \frac{1}{N}\sum_{n=1}^{N} p^{\text{old}}(\omega_l \mid x_n, \Theta_l^{\text{old}}) \tag{11.17}$$

The m-step involves maximizing the auxiliary function with respect to the parameters $\{\Theta_l\}_{l=1}^{L}$. The auxiliary function can be written as

$$Q(\Theta^{\text{new}}, \Theta^{\text{old}}) = \sum_{n=1}^{N}\sum_{l=1}^{L} p^{\text{old}}(\omega_l \mid x_n, \Theta_l^{\text{old}}) \log P^{\text{new}}(\omega_l) p^{\text{new}}(x_n \mid \omega_l, \theta_l^{\text{new}}) \tag{11.18}$$

where

$$p^{\text{new}}(x_n \mid \omega_l, \Theta_l^{\text{new}}) = \sum_{m=1}^{M} p^{\text{new}}(m_l) p^{\text{new}}(x_n \mid m_l, \Theta_{ml}^{\text{new}}) \tag{11.19}$$

Writing $\gamma_{nl} = p^{\text{old}}(\omega_l \mid x_n, \Theta_l^{\text{old}})$, the auxiliary function can be written as the sum of L auxiliary functions, one for each mixture model:

$$Q(\Theta^{\text{new}}, \Theta^{\text{old}}) = \sum_{n=1}^{N}\sum_{l=1}^{L} \gamma_{nl} \log P^{\text{new}}(\omega_l) p^{\text{new}}(x_n \mid \omega_l, \theta_l^{\text{new}}) \tag{11.20}$$

$$Q(\Theta^{\text{new}}, \Theta^{\text{old}}) = \sum_{l=1}^{L} \hat{Q}_l(\Theta^{\text{new}}, \Theta^{\text{old}}) \tag{11.21}$$

$$\text{where} \quad \gamma_{nl} = p(\omega_l \mid x_n, \Theta_l) = \frac{p(x_n \mid \omega_l, \theta_l) P(\omega_l)}{\sum_{j=1}^{L} p(x_n \mid \omega_j, \theta_j) P(\omega_j)} \tag{11.22}$$

$$\text{and} \quad \hat{Q}_l(\Theta_l^{\text{new}}, \Theta_l^{\text{old}}) = \sum_{n=1}^{N} \gamma_{nl} \log P^{\text{new}}(\omega_l) p^{\text{new}}(x_n \mid \omega_l, \theta_l^{\text{new}}) \tag{11.23}$$

The procedure for maximising the overall likelihood of a WGMM is outlined in Algorithm 1. It consists of an outer EM loop, which are nested in L inner EM loops. Each time the outer loop is traversed, the mixing weights $p(\omega_l)$ are updated according to Eq. (11.17) and the L inner loops are iterated to update the mixing weights $pl(m)$, means μ_{lm}, and covariances Σ_{lm} for each of the components. It should be noted that it is not necessary to iterate the inner loops to converge on each outer EM step, since it is only necessary to increase the

auxiliary function to ensure convergence of the overall likelihood to a local maximum.

Algorithm 1: **WGMM Algorithm**

1. Make an initial estimate of all GMM parameter values $\{\Theta_l\}_{l=1}^{L}$ and $p(\omega_l)$.

2. **Iterate** outer *E*-step and outer *M*-step until the change in auxiliary function (Eq. 11.18) between iterations is less than some convergence threshold $\text{WGMM}_{\text{converge}}$.

3. **Outer EM E-step**:

   (a) Compute $\gamma_{nl} = p^{\text{old}}(\omega_l \mid x_n, \Theta_l^{\text{old}})$.

   (b) Evaluate an auxiliary function $Q(\Theta^{\text{new}}, \Theta^{\text{old}})$ as in Eq. (11.18).

4. **Outer EM** M-step:

   (a) **Inner EM steps**

   For each GMM modeling the class-conditional *pdf* of classes $\omega_l = \{1, \ldots, L\}$ do (update the parameter values of each individual GMM using the GMM-EM algorithm until convergence).

   (b) Find new values for the $WGMM$ mixing coefficients, Θ^{new}, that maximizes the auxiliary function given in step 3(b) above.

5. Iterate steps 2–4 until the convergence criteria are satisfied.

Finally, we combine our WGMM model with MRF in the same manner as Zhang *et al.* [24] combined *GMM* with MRF. The $WGMM^{MRF}$ model is based on Eq. (11.15) except that the mixing coefficients $p(\omega_l)$ are replaced with a MRF-MAP estimate $p(y_n = \omega_l \mid \aleph_n)$ using ICM algorithm [29]. The auxiliary function given in Eq. (11.18) is rewritten to include the MRF hidden model as follows:

$$Q(\Theta^{\text{new}}, \Theta^{\text{old}}) = \sum_{n=1}^{N} \sum_{l=1}^{L} p^{\text{old}}(\omega_l \mid x_n, \Theta_l^{\text{old}}) \log(p(y_n = \omega_l \mid \aleph_n) p^{\text{new}}(x_n \mid \omega_l, \Theta_l^{\text{new}}) \tag{11.24}$$

The update equations for the mean and covariances in the GMM-EM algorithm remain unchanged. The MRF-MAP estimate is combined in the conditional density function $p^{\text{old}}(\omega_l \mid x_n, \theta_l^{\text{old}})$ as

$$\gamma_{nl} = p(\omega_l \mid x_n, \theta_l) = \frac{p(x_n \mid \omega_l, \Theta_l)p(y_n = \omega_l \mid \aleph_n)}{\sum_{j=1}^{M} p(x_n \mid \omega_j, \Theta_j)p(y_n = \omega_j \mid \aleph_n)} \tag{11.25}$$

The WGMM$^{\text{MRF}}$-EM algorithm is used to determine the ML estimates of the parameter values by iterating the WGMM-EM algorithm while constraining the density estimation with the hidden MRF model. For supervised learning, the labelled training data is used for the initialization of the WGMM and WGMM$^{\text{MRF}}$ models, to give us WGMM$_{\text{S}}$, and WGMM$_{\text{S}}^{\text{MRF}}$ and no training data is used for the unsupervised learning case, WGMM$_{\text{U}}$, and WGMM$_{\text{U}}^{\text{MRF}}$.

11.4.4 Combination of Image Segmentation Experts

In the previous section we developed four new models of image segmentation and mentioned the use of different experts based on different texture features that rely on them. It is beneficial to fuse the decisions of different experts on a per pixel basis. In this section we detail the conventionally used strategy of classifier decision combination, called "ensemble based combination rules," and then propose a novel strategy for combining expert outputs, called "adaptive weighted model (AWM)." First of all, we describe a generic framework of combination, and then discuss the combination strategies within that framework.

11.4.4.1 Expert Combination Framework and Nomenclature

The image to be segmented can be represented as a 1-D array $X = \{x_1, \ldots, x_N)$, where x_n is an input feature for the pixel n and N is the total number of pixels in the image. Let the estimate of the segmentation be denoted by array $\hat{Y} = \{\hat{y}_1, \ldots, y_N)$. It is assumed that the number of classes is predetermined from a set of known class labels $\omega_l \in \{1, \ldots, L\}$, and therefore, the estimated class label of pixel n is indicated as $\hat{y}_n = \omega_l$.

We assume that there are R image segmentation experts, where the rth expert provides a segmentation decision for a given pixel feature x_n from a set of learnt parameter vectors θ_r. Using a WGMM expert, the parameter vector θ_r of each expert is defined as a set of component mixing coefficients $pl(m)$, means μ_{lm}, and

covariances Σ_{lm} from each of the M component Gaussians, $m \in \{1, \ldots, M\}$, for each class $\omega_l \in \{1, \ldots, L\}$. On segmentation of an image, the rth expert provides an estimate of the *a posteriori* probability of a feature vector associated with a pixel x_n, belonging to a given class ω_l as $p(\hat{y}_n = \omega_l \mid x_n, \theta_r)$, for $\forall_n = (1, \ldots, N)$. In order to combine the decisions of different experts, the joint probability of all segmentation decisions is required. Using the Bayes rule, the combined *a posteriori* probability can be computed from the segmentation experts for class ω_l as follows:

$$p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) = \frac{p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) p(\omega_l)}{p(x_n, \theta_1, \ldots, \theta_R)} \tag{11.26}$$

where $p(\omega_l)$ is the prior probability (assumed to be set equally for all classes as $1/R$) for each class ω_l, and $p(x_n, \theta_l, \ldots, \theta_R)$ is the unconditional joint probability defined as

$$p(x_n, \theta_1, \ldots, \theta_R) = \sum_{k=1}^{L} p(\hat{y} = \omega_k \mid x_n, \theta_1, \ldots, \theta_R) p(\omega_k) \tag{11.27}$$

On the basis of this nomenclature and equal priors from each class, in the following two sections we detail the "ensemble-based combination rules" (section 4.4.2), and then propose a novel strategy for combining results, called "adaptive weighted model (AWM)" (section 4.4.3)

11.4.4.2 Ensemble-Based Combination Rules

Kittler [33] proposed a set of very popular rules for combining probability outputs from a number of experts. These rules are stated as follows:

Product

$$(Prod) \qquad p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) = \frac{\prod_{r=1}^{R} p(\hat{y} = \omega_l \mid x_n, \theta_r)}{\sum_{j=1}^{L} \prod_{r=1}^{R} p(\hat{y} = \omega_j \mid x_n, \theta_r)}.$$

Sum

$$(Sum) \qquad p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) = \frac{1}{R} \sum\nolimits_{r=1}^{R} p(\hat{y} = \omega_l \mid x_n, \theta_r)$$

Max

$$(Max) \qquad p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) = \max\nolimits_{r=1}^{R} (\hat{y} = \omega_l \mid x_n, \theta_r)$$

Min

$$(Min) \qquad p(\hat{y} = \omega_l \mid x_n, \theta_1 \ldots \theta_R) = \min\nolimits_{r=1}^{R} (\hat{y} = \omega_l \mid x_n, \theta_r)$$

Majority Voting (Mv)

$$p(\hat{y} = \omega_l \,|\, x_n, \theta_1 \ldots \theta_R) = \frac{\sum_{r=1}^{R} \Delta_{lr}}{R}$$

$$\text{where} \quad \Delta_{lr} = \begin{cases} 1 & \text{if } p(\hat{y} = \omega_l \,|\, x_n, \theta_1 \ldots \theta_r) \\ & \quad = \max_{r=1}^{R} p(\hat{y} = \omega_l \,|\, x_n, \theta_j) \\ 0 & \text{otherwise} \end{cases}$$

The above combination rules have been used in several studies and form the basis of our baseline comparison.

11.4.4.3 Average Weighted Model (AWM) Classifier Combination

In our proposed approach, the expert decisions are modeled as a probability density function. From a linear opinion pool of R experts, assume that the rth segmentation expert provides an estimate of the *a posteriori* probability.

$$p(\hat{y}_n \,|\, r, x_n) = p(\hat{y}_n \,|\, x_n, \theta_r) \quad \forall n = (1, \ldots, N) \tag{11.28}$$

We assume that accompanying this pdf is a linear weight or mixing coefficient, $p(r)$, indicating the contribution of the rth expert in the joint pdf, $p(\hat{y} \,|\, x, \Theta)$, resulting from the combination of experts. The vector Θ is the complete set of parameters describing the combined pdf. Hence, following the expert combination, the complete pdf can be written as

$$p(\hat{y}_n \,|\, x_n, \Theta) = \sum_{r=1}^{R} p(r) p(\hat{y}_n \,|\, r, x_n) \tag{11.29}$$

given that the mixing coefficients satisfy the following constraints: $\sum_{r=1}^{R} p(r) = 1$ and $0 \leq p(r) \leq 1$. If we treat the weighted contribution of each expert in the unconditional distribution as probabilities, then statistical models such as mixture of experts (MOE) framework [34] can be trained to learn the individual classifier and weight contribution distributions. For this we propose using the GMM using EM algorithm. We now present a method for identifying the weights in a probabilistic manner motivated by the MOE framework. Our proposed approach is, however, different to the conventional MOE method in two ways: (i) First, the *a posteriori pdf* from each segmentation expert remains fixed having been generated during segmentation; (ii) second, the mixing coefficients for

each expert, $p(r)$, are determined in an unsupervised manner through statistical methods.

11.4.4.3.1 Maximum Likelihood Solution. The mixing coefficient parameter values for each expert can be determined using the ML principle by forming a likelihood function. Assume that we have the complete dataset, ψ, of combined decisions from segmentation experts for each data point, where $\psi = \{\hat{y}_1, \ldots, \hat{y}_N)$, and it is drawn independently from the complete distribution $p(\hat{y} \mid x, \Theta)$. Then the joint occurrence of the whole dataset is given as

$$p(\psi \mid \Theta) = \prod_{n=1}^{N} \sum_{r=1}^{R} p(r) p(\hat{y}_n \mid r, x_n) \equiv \zeta(\Theta) \tag{11.30}$$

For simplicity, the above likelihood function can be rewritten and expressed as a log likelihood as follows:

$$\log \zeta(\Theta) = \sum_{n=1}^{N} \log p(\hat{y}_n \mid \Theta) \equiv \sum_{n=1}^{N} \log \sum_{r=1}^{R} p(r) p(\hat{y}_n \mid r, x_n) \tag{11.31}$$

For the above equation, it is not possible to find the ML estimate of the parameter values Θ directly because of the inability to solve $\frac{\partial \zeta}{\partial \Theta} = 0$ [23]. Our approach used to maximising the likelihood $\log \zeta(\Theta)$ is based on the EM algorithm proposed in the context of missing data estimation [35].

11.4.4.3.2 AWM Parameter Estimation Using EM Algorithm. The EM algorithm attempts to maximize an estimate of the log likelihood that expresses the expected value of the complete data log likelihood conditional on the data points. By evaluating an auxiliary function, Q in the E-step, an estimate of the log likelihood can be iteratively maximized using a set of update equations in the M-step. Using the AWM likelihood function from Eq.(11.30) the auxiliary function for the *AWM* is defined as

$$Q(\Theta^{\text{new}}, \Theta^{\text{old}}) = \sum_{n=1}^{N} \sum_{r=1}^{R} p^{\text{old}}(r \mid \hat{y}_n) \log(p^{\text{new}}(r) p(\hat{y}_n \mid r, x_n)) \tag{11.32}$$

It should be noted that the *a posteriori* estimate $p(\hat{y}_n \mid r, x_n)$ for the nth data point from the rth segmentation expert remains fixed. The conditional density function $p^{\text{old}}(r \mid \hat{y}_n)$ is computed using the Bayes rule as

$$p^{\text{old}}(r \mid \hat{y}_n) = \frac{p(\hat{y}_n \mid r, x_n) p(r)}{\sum_{j=1}^{R} p(\hat{y}_n \mid j, x_n) p(j)} \tag{11.33}$$

In order to maximize the estimate of the likelihood function given by the auxiliary function, update equations are required for the mixing coefficients. These can be obtained by differentiating with respect to the parameters set equal to zero. For the *AWM*, the update equations are taken from [27]. For the rth segmentation expert

$$p^{\text{new}}(r) = \frac{1}{N} \sum_{n=1}^{N} p^{\text{old}}(r \mid \hat{y}_n) \tag{11.34}$$

The complete AWM algorithm is shown below.

Algorithm 2: **AWM Algorithm**

1. **Initialise**: Set $p(r) = 1/R$.
2. **Iterate**: Perform E-step and M-step until the change in Q function, Eq. (11.31), between iterations is less than some convergence threshold $AVM_{converge} = 25$.
3. **EM E-step**:

   (a) Compute $p^{\text{old}}(r \mid \hat{y}_n)$ using Eq. (11.32).

   (b) Evaluate the Q function, the expectation of the log-likelihood of the complete training data samples given the observation, x_n, and the current estimate of the parameters using Eq. (11.31).

4. **EM M-step**: This consists of maximising $\hat{Q}$ with respect to each parameter in turn:

   1. The new estimate of the segmentation expert weightings for the rth component $P^{\text{new}}(r)$ is given by Eq.(11.33).

11.4.4.3.3 Estimating the *A Posteriori* Probability. Using the *AWM* combination strategy in mammographic CAD, a *posteriori* estimates are required for each data point following the experts' combination (one for the normal and one for the suspicious class). To determine these estimates, the AWM model is computed for the first class, thereby obtaining the *a posteriori* estimate $p(\hat{y}_n = \omega_1 \mid x_n, \Theta)$. From this, the estimate of the second class is determined as $p(\hat{y}_n = \omega_2 \mid x_n, \Theta) = 1 - p(\hat{y}_n = \omega_1 \mid x_n, \Theta)$. We now proceed to the results

section to evaluate our novel contributions of weighted GMM segmentation experts and the novel AWM combination strategy.

11.4.5 Results of Applying Image Segmentation Expert Combination

The aim of our experiments was to (i) perform a comparison between the four proposed models of image segmentation. The baseline comparison with a simple GMM based image segmentation and an MRF model in [18] shows that our proposed models easily outperform the baseline models. (ii) To compare the performance of the *AWM* combination strategy against the ensemble combination rules. Section 11.4.5.1 compares the four models on the two databases, and section 11.4.5.2 compares the AWM approach with ensemble combination rules approach on the two databases.

Our segmentation performance evaluation is performed on 400 mammograms selected from the DDSM. The first 200 mammograms contain lesions and the remaining 200 mammograms are normal (used only for training purposes). Each of these mammograms has been categorized into one of the four groups representing different breast density, such that each category has 100 mammograms. The partitioning of the mammograms has been performed manually on the basis of the target breast density according to DDSM ground truth. The results will be reported in terms of the A_z value that represents the area under the ROC curve as well as sensitivity (the segmentation evaluation for testing is based on ground-truth information as given in DDSM).

The grouping of mammograms by breast density is applicable only to the supervised approaches. Supervised approaches segmenting a mammogram with a specific breast density type use a trained observed intensity model constructed with only training samples from that breast type. Thus, each trained observed intensity model will be specialized in the segmentation of a mammogram with a specific breast type. We adopt a fivefold cross-validation strategy. Using this procedure, a total of five training and testing trials are conducted, and each time the data appearing in training does not appear as testing. For each of the fivefolds, equal numbers of normal and suspicious pixels are used to represent training examples from their respective classes. These sample pixels are randomly sampled from the training images. In the unsupervised

Table 11.10: Mean A_Z for each breast type and segmentation strategy.

Breast type	$WGMM_S$	$WGMM_S^{MRF}$	$WGMM_U$	$WGMM_U^{MRF}$
1	0.68	**0.70**	0.66	0.59
2	**0.66**	**0.66**	0.**66**	0.60
3	0.72	**0.80**	0.75	0.75
4	0.66	**0.76**	0.68	0.74
Mean	0.68	**0.73**	0.68	0.67

Winning strategies are given in bold.

case, there is no concept of training and testing and each image is treated individually.

11.4.5.1 Comparison of the Four models ($WGMM_S$, $WGMM_U$, $WGMM_S^{MRF}$, and $WGMM_U^{MRF}$)

A cross-validation approach is used to determine the optimal number of component Gaussians, for each breast type. The determined value of m is then used for all training folds comprising each breast type. To determine the optimal value of m, models with a different number of components are trained and evaluated with a $WGMM_S$ strategy, using an independent validations set. Model fitness is quantified by examining the log likelihood resulting from the validation set. Training files are created by taking 200 samples randomly drawn with replacement from each normal and abnormal images for each breast type. For training we use 50 training images per breast type ($n = 25$ normal, $n = 25$ abnormal) giving a training size of 10,000 samples per breast type. Repeating the procedure for 50 remaining validation image per breast type, we get 10,000 samples for validation.

In our evaluation procedure the aim is to determine the correct number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) in order to plot the ROC curve. A detailed summary of how each segmented region is classed as one of these is detailed in [18]. The results are shown in Table 11.10 grouped on the basis of breast density. It is easily concluded that the supervised strategy with MRF is a clear winner. Interestingly, the performance of this method is superior for denser images compared to fatty ones. A simple

explanation for this phenomenon could be based on the model order selection where $m = 1$ for the abnormal class of the fatty breast types. A more sophisticated approach to determining model order might improve the segmentation of these breast types. Without the hidden MRF model, the supervised strategy is inferior to the unsupervised approach on the denser breasts.

11.4.5.2 Comparison of Combination Strategies: Ensemble Combination Rules vs. AWM

In order to develop a number of experts that can be combined, we extract different gray-scale and texture data per pixel in the images. The gray-scale values of the pixels are intensity values, and texture features are extracted from pixel neighborhood. The following table shows the different feature experts used in our analysis based on different features. Each expert can be implemented with one of the four segmentation models described earlier.

Expert	Description of pixel feature space	Dimensionality
gray	Original gray scale	1
enh	Contrast enhanced gray scales	1
dwt^1	Wavelet coefficients from $\{D^1_{LH}, D^1_{HH}, D^1_{HL}, S^1_{LL}\}$	4
dwt^2	Wavelet coefficients from $\{D^2_{LH}, D^2_{HH}, D^2_{HL}, S^2_{LL}\}$	4
dwt^3	Wavelet coefficients from $\{D^3_{LH}, D^3_{HH}, D^3_{HL}, S^3_{LL}\}$	4
$laws^1$	Laws coefficients from $E5$ impulse response matrix	5
$laws^2$	Laws coefficients from $L5$ impulse response matrix	5
$laws^3$	Laws coefficients from $R5$ impulse response matrix	5
$laws^4$	Laws coefficients from $W5$ impulse response matrix	5
$laws^5$	Laws coefficients from $S5$ impulse response matrix	5

We now present the results on 200 test mammograms that contain lesions. The details of training and testing scheme are the same as detailed in section 11.4.2. As we mentioned earlier, each breast is classified as one of the four types (1, predominantly dense; 2, fat with fibroglandular tissue; 3, heterogeneously dense; and 4, extremely dense) and the results are presented for data from each type. Table 11.11 shows the test results on sensitivity of the

Table 11.11: Mean sensitivity for each testing strategy for DDSM image database

	Breast type 1	Breast type 2	Breast type 3	Breast type 4
$WGMM_S$	$laws^1$	$laws^4$	$laws^4$	$laws^4$
	0.740	0.545	0.675	0.510
$WGMM_S^{MRF}$	$laws^1$	$laws^1$	enh	$laws^1$
	0.690	**0.650**	0.650	**0.640**
$WGMM_U$	enh	$laws^2$	enh	$laws^1$
	0.525	0.575	0.660	0.550
$WGMM_U^{MRF}$	$laws^1$	$laws^1$	$laws^4$	$laws^1$
	0.690	0.640	**0.690**	0.540

Results are shown for all breast types. Winning segmentation expert are shown in bold per breast type.

different segmentation models with different features without expert combination. The following key conclusion can be drawn from these results: (a) A single feature is not always the winning feature. In general, features enh, $laws_1$, and $laws_4$ do quite well. (b) It is easier to segment fatty breasts as opposed to dense breasts which is to be expected. (c) Models using MRF work better than those that do not use them. (d) There is no clear cut winner between supervised and unsupervised strategy; depending on which features they use, they can outperform the other. (e) For three of the breast types 1, 2, and 4, the model $WGMM_S^{MRF}$ is a clear winner, whereas for breast type 3, $WGMM_U^{MRF}$ performs the best.

Table 11.12: Mean sensitivity for each combination strategy for DDSM database

	Breast type 1	Breast type 2	Breast type 3	Breast type 4
$WGMM_S$	M_v	AWM	AWM	Min
	0.510	0.520	0.701	0.505
$WGMM_S^{MRF}$	AWM	Sum	AWM	AWM
	0.575	0.630	**0.727**	**0.680**
$WGMM_U$	AWM	M_v	M_v	M_v
	0.320	0.532	0.515	0.525
$WGMM_U^{MRF}$	AWM	AWM	AWM	AWM
	0.550	**0.667**	0.705	0.625

Results are shown for all breast types. Winning combination method shown in bold per breast type.

Table 11.13: The results from best performing (a) expert strategy and (b) AWM combination strategy.

	T	*Seg*	*Expert*	*Sens*	*% mass*
(a)	1	$WGMM_S$	$laws^1$	**0.740**	.15
	2	$WGMM_S^{MRF}$	$laws^1$	0.650	.23
	3	$WGMM_U^{MRF}$	$laws^1$	0.690	.31
	4	$WGMM_S^{MRF}$	$laws^1$	0.640	.28
	T	*Seg*	*Cmb*	*Sens*	*% mass*
(b)	1	$WGMM_S^{MRF}$	*AWM*	0.575	.25
	2	$WGMM_U^{MRF}$	*AWM*	**0.667**	.26
	3	$WGMM_S^{MRF}$	*AWM*	**0.727**	.38
	4	$WGMM_S^{MRF}$	*AWM*	**0.680**	.37

Winning strategy shown in bold. T = breast type; *Seg* = segmentation strategy; *Cmb* = combination strategy; *Sens* = sensitivity; % *mass* = mean percentage of target lesion detected as true positive.

We next compare the ensemble combination rules with the AWM expert combination strategy on the four breast type data testing. The results are shown in Table 11.12. The key results can be summarized as follows: (a) The AWM method result always turns out to be the overall best result compared to all ensemble combination rules on all breast types. (b) The AWM results are best with the $WGMM_S^{MRF}$ segmentation method on breast types 1, 3, and 4, and best with WGMM_U^{MRF} on breast type 2. (c) The combination methods Max and $Prod$ never win. (d) Segmentation models using MRF are better than those that do not use them.

In Table 11.13 we compare single best experts with the best combination of experts for the four breast types. The results show that only on breast type 1, using the single best expert $WGMM_S$ with $laws_1$, features will outperform all other experts and combination of experts (sensitivity of 0.74). For the remaining three breast types, the *AWM* expert combination method is the best. For breast types 3 and 4 (dense breasts), the supervised learning based models with *MRF* are better, whereas for fatty breast of type 2, unsupervised learning model with MRF is the best.

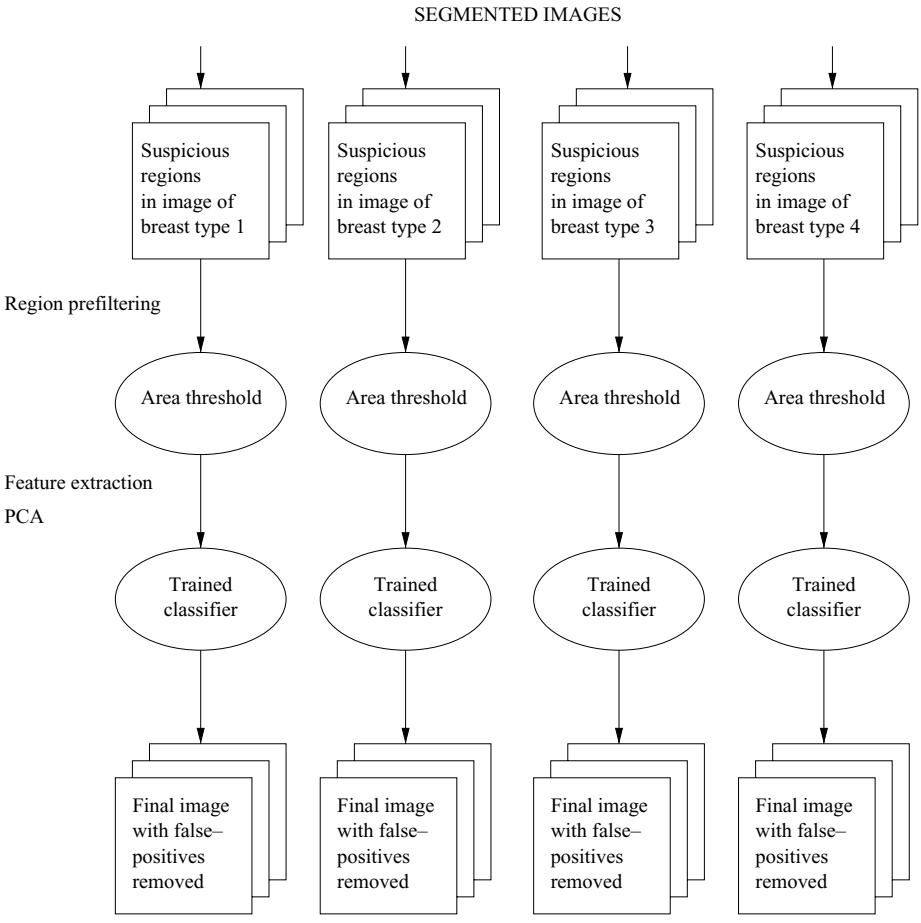


Figure 11.7: Schematic overview of false-positive reduction strategy within the adaptive knowledge-based model.

11.5 A Framework for the Reduction of False-Positive Regions

This section describes the approach used within the adaptive knowledge-based model for the reduction of false-positive regions. Figure 11.7 shows a schematic overview of the approach adopted. Using the actual breast type grouping predicted by the breast classification component, a segmented mammogram is directed to one of four process flows. Each process flow, shown in Fig. 11.7,

comprises the same functionality. This is discussed in more detail in the following subsections.

11.5.1 Postprocessing Steps for Filtering Out False Positives

11.5.1.1 Region Prefiltering

Feature extraction is computationally expensive. A common strategy [6, 7, 36] to reduce the number of regions considered for false-positive reduction is achieved by applying a size test. By eliminating suspicious regions smaller than a predefined threshold T_{area}, the number of false-positive regions can be reduced. For the expert radiologist interpreting a film mammogram during screening, it is common to disregard any suspicious ROI less than 8 mm in diameter [37]. In mammographic CAD with computer automation, the size threshold is reduced and a common value for T_{area} is the number of pixels corresponding to an area of 16 mm^2 [6, 7, 36]. In the adaptive knowledge-based model, the area threshold is set at 19.5 mm^2 corresponding to a region diameter of 5 mm for all breast type groupings. The DDSM used in this evaluation are digitized such that each pixel is 50 μm. Following subsampling by a factor of four, an area threshold of 19.5 mm^2 is equivalent to $T_{\text{area}} = 122$ pixels, thus any suspicious region following segmentation with an area less than this value is marked as normal.

11.5.1.2 Feature Extraction

Features are extracted to characterise a segmented region in the mammogram. Feature vectors from masses are assumed to be considered different from normal tissue, and based on a collection of their examples from several subjects, a system can be trained to differentiate between them. The main aim is that features should be sensitive and accurate for reducing false positives. Typically a set or vector of features is extracted for a given segmented region.

From the pixels that comprise each suspicious *ROI* passing the prefiltering size test described above, a subset of gray scale, textural, and morphological features used in previous mammographic studies are extracted. The features extracted are summarized in Table 11.14.

Table 11.14: Summary of features extracted by feature grouping giving 316 features in total

Grouping	Type	Description	Number
Gray scale	Histogram	Mean, variance, skewness, kurtosis, and entropy.	5
Textural	SGLD	From SGLD matrices constructed in 5 different directions and 3 different distances 15 features [38, 39] are extracted.	15×15
	Laws	Texture energy [6] extracted from 25 mask convolutions.	5×5
	DWT	From DWT coefficients of 4 subbands at 3 scales the following statistical features are extracted: mean, standard deviation, skewness, kurtosis.	4×12
	Fourier	Spectral energy from 10 Fourier rings.	10
	Fractal	Fractal dimension feature.	1
Morphological	Region	Circularity [4] area.	2

11.5.1.3 Principal Component Analysis

The result of feature extraction is a 316-dimensional feature vector describing various gray-scale histogram, textural, and morphological characteristics of each region. The curse of dimensionality [27] is a serious constraint in many pattern recognition problems and to maintain classification performance, the dimensionality of the input feature space must be kept to a minimum. This is especially important when using an ANN classifier, to maintain a desired level of generalization [32]. Principal component analysis (PCA) is a technique to map data from a high-dimensional space into a lower one and is used here for such a purpose.

To use PCA in the adaptive knowledge-based model in an unbiased way, the PCA coefficients, comprising eigenvalues and eigenvectors, are determined from an independent training set. In mapping to a lower dimensionality, only eigenvalues ≥ 1.0 are considered and the eigenvectors from training are applied to a testing pattern. Testing and training folds are formed using 10-fold cross validation [32] such that an unbiased PCA transformation can be obtained for each testing sample.

11.5.1.4 Artificial Neural Network Classification

Using a labelled training set, an ANN classifier can be trained using supervised learning algorithms to discriminate between normal and abnormal regions. Features from representative training samples are provided during supervised learning and the weights of the ANN are updated until the generalization ability of the classifier starts to decrease measured on a separate validation set. Implementation in the adaptive knowledge-based model results in the construction of a separate ANN classifier for each breast type grouping. Only regions from mammograms of the same mammogram type will be considered for each ANN. Each ANN is a three-layer feed-forward network comprised of a different number of hidden nodes and two output nodes (normal, abnormal). The optimal number of hidden nodes is determined for each ANN individually. To ensure an unbiased result and that every sample is used at least once in training and testing, a 10-fold cross validation strategy [32] is employed. No sample appears simultaneously in training and test. Additionally a validation set is used (comprising 10% of the training samples) to prevent over-fitting of the ANN to the training set. The feed-forward ANN is trained using a back-propagation with momentum learning function (learning rate $\eta = 0.01$, momentum $\mu = 0.5$) together with a softmax activation function and used on test to give an estimate of the *a posteriori* probability of each pattern for each class.

11.5.2 Results from DDSM Abnormal Images

This section gives the results obtained from applying the false-positive reduction strategy to 200 segmented DDSM mammograms containing breast lesions as defined by an expert radiologist. Each mammogram has been assigned to one of four breast type groupings such that 50 mammograms exist for each grouping. Quantitative measures of performance are given in terms of sensitivity over all mammograms in each breast group together with the average number of false-positive regions per image.

11.5.2.1 Feature Extraction and PCA

Following feature extraction of the 316-dimensional feature vector for each sample, PCA is used to map the sample data from a higher dimension to that of

Table 11.15: Sensitivity and average number of false-positive regions per image over all 50 abnormal mammograms

	After segmentation		After segmentation prefiltering		After false-positive reduction	
Breast type	Sensitivity	*FP/i*	Sensitivity	*FP/i*	Sensitivity	*FP/i*
1	0.57	163.31	0.51	9.31	0.40	3.26
2	0.58	132.09	0.56	8.26	0.48	4.40
3	0.72	132.55	0.70	6.90	0.66	4.14
4	0.66	158.79	0.64	10.27	0.60	3.56
Mean	0.63	146.69	0.60	8.68	0.54	3.84

After expert segmentation with $WGMM_S^{MRF}$ combined using *AWM*; after region prefiltering using $T_{\text{area}} = 122$; after false-positive (*FP*) reduction using classifier operating point, by breast type.

a lower one. Using the unbiased PCA strategy described above only eigenvalues ≥ 1.0 are considered, resulting in a 37-dimensional feature vector.

11.5.2.2 Optimization of Networks

To optimize the number of hidden nodes, using 10-fold cross validation, different ANN models are evaluated. For the evaluated ANN model, performance in discriminating between abnormal and normal regions is determined using receiver operating characteristic (ROC) analysis [40]. By calculating the area under the ROC curve (A_Z), a quantitative measure of performance can be determined.

11.5.2.3 Results from False-Positive Reduction

Table 11.15 summarizes the results from applying the false-positive reduction strategy to 200 abnormal segmented DDSM mammograms. Three sets of results are shown for each stage in the false-positive approach described for each breast type grouping.

The first column shows the sensitivity and average number of false positives per image following mammogram segmentation. The segmentation was obtained by combining 10 segmentation expert outcomes using the *AWM* described earlier. Each expert was constructed using the *WGMM* constrained with a *MRF* utilizing a supervised learning approach $WGMM_S^{MRF}$.

The second column shows the sensitivity and average number of false-positives regions per image obtained after applying the region prefiltering. These results demonstrate the utility of the region prefiltering stage. The average number of false-positive regions per image has dropped from approximately 147 to just 9 when testing on the complete dataset of 200 abnormal mammograms. This result has been obtained at a reduction in the sensitivity to the detection of breast lesions, from 0.63 to 0.60, for all breast types.

The final column shows the results obtained after classifying each region passing the prefiltering using an optimized ANN based on the 37-dimensional PCA feature vector for each sample. Using ROC analysis, the threshold for the detection of positive cases is set using the operating point of each ANN [40]. From these results it can be seen that the sensitivity is reduced still further to just 0.54 for all 200 abnormal mammograms, with a reduced average number of false-positive regions per image of 3.84. The results indicate that the biggest drop in sensitivity is obtained for the fatty breasts, breast types 1 and 2. This may be attributed to the increased variability of breast lesions in these breast types compared with that of the denser breasts.

11.5.3 Key Observations

The above discussion has described an approach to the reduction of false-positive region from segmented images containing suspicious ROI. The following key observations can be drawn:

1. *Region prefiltering*: prefiltering regions based on their area is a quick and simple method to reducing false-positive regions while maintaining similar levels of sensitivity prior to filtering. The area threshold T_{area} is defined for a circular region with a diameter of 5 mm. This is a similar value to that used in other studies and is stricter than that used by expert radiologists when interpreting film screen mammograms.

2. *Feature extraction*: By surveying previous studies, a subset of features for use in the reduction of false-positives regions has been evaluated. These features capture, morphological, gray scale, and texture information about each region. Using an unbiased implementation of PCA, the 316-dimensional feature space is reduced to a 37-dimensional feature space.

3. *Sensitivity in the detection of breast lesions*: Following evaluation of the FP reduction strategy on 200 abnormal segmented abnormal DDSM mammograms, sensitivity levels dropped by over 8% but the average number of false-positive regions per image drops by approximately 98%. Varying the threshold on the ANN classifier using ROC analysis, the expert radiologist can select a threshold that varies the available sensitivity at the expense of an acceptable number of false-positive regions.

11.6 Evaluation of the Knowledge-Based Model

This section evaluates the performance of a given configuration of the adaptive knowledge-based model in predicting the optimal pipeline of image processing operators used for the CAD of breast cancer. This performance is compared to that obtained by keeping the pipeline fixed. Contrast enhancement and image segmentation are the key components in a mammographic CAD system. For these key components, sections 11.3 and 11.4, respectively, have demonstrated that a knowledge-based framework is superior to the single best method in each case. Parameterized versions of these components have been engineered for individual mammogram groupings. These groupings are based on the mammographic breast density and a mechanism for its prediction. Evaluation of the performance of each parameterized version of the knowledge-based component presented in the previous sections has been performed using the target mammogram breast grouping. In this section, the complete adaptive knowledge-based model is evaluated using the predicted breast group.

Section 11.6.1 evaluates the knowledge-based contrast enhancement and segmentation components using the predicted breast type grouping using 200 abnormal mammograms from the DDSM. Following this, section 11.6.2 evaluates the complete adaptive knowledge-based model using a dataset of 400 mammograms. This dataset comprises 200 normal and 200 abnormal mammograms comprising 50 images of each type from each of the four breast types. Results for segmentation and following false-positive reduction are presented. Finally section 11.6.3 presents key observations.

11.6.1 Expert Contrast Enhancement and Segmentation of Abnormal Images with Adaptive Knowledge-Based Model

11.6.1.1 Dataset and Adaptive Knowledge-Based Model Configuration

This section presents the results from evaluating the optimal contrast enhancement and segmentation knowledge-based components of the adaptive knowledge-based model on a dataset of 200 DDSM mammograms containing abnormalities. The 200 mammograms comprise 50 images from each of four different breast types. To obtain a testing result for each mammogram, knowledge-based components utilize separate training and testing folds such that no image from a test fold exists in a corresponding training fold. Training data for the abnormal mammograms is based on redefined DDSM ground truth boundaries.

Figure 11.8 shows the configuration of the adaptive knowledge-based model for contrast enhancement and mammogram segmentation used for performance evaluation. Enhancement and segmentation experts are identified in the black boxes. Knowledge-based components, providing optimal enhancement and optimal segmentation, are identified in dotted boxes. Associated with each expert and knowledge-based component in Fig. 11.8 is a table with four rows, one for each breast type. The right-hand column of the table identifies the performance of the associated expert or knowledge-based component for all mammograms of the predicted breast type. This performance measure is computed differently for contrast enhancement and segmentation components as follows:

(a) *Enhancement component*: Performance is measured by the mean percentage improvement in the segmentation of the contrast-enhanced image compared to that of the unenhanced original for all mammograms of a given breast type.

(b) *Segmentation component*: Performance is measured by the mean area (A_Z) under the ROC curve, for all mammograms of a given breast type. Use of this measure in evaluating the adaptive knowledge-based model reflects the underlying sensitivity and false-positive count across all ROC thresholds and has been used in other studies [41] to compare classification tasks.

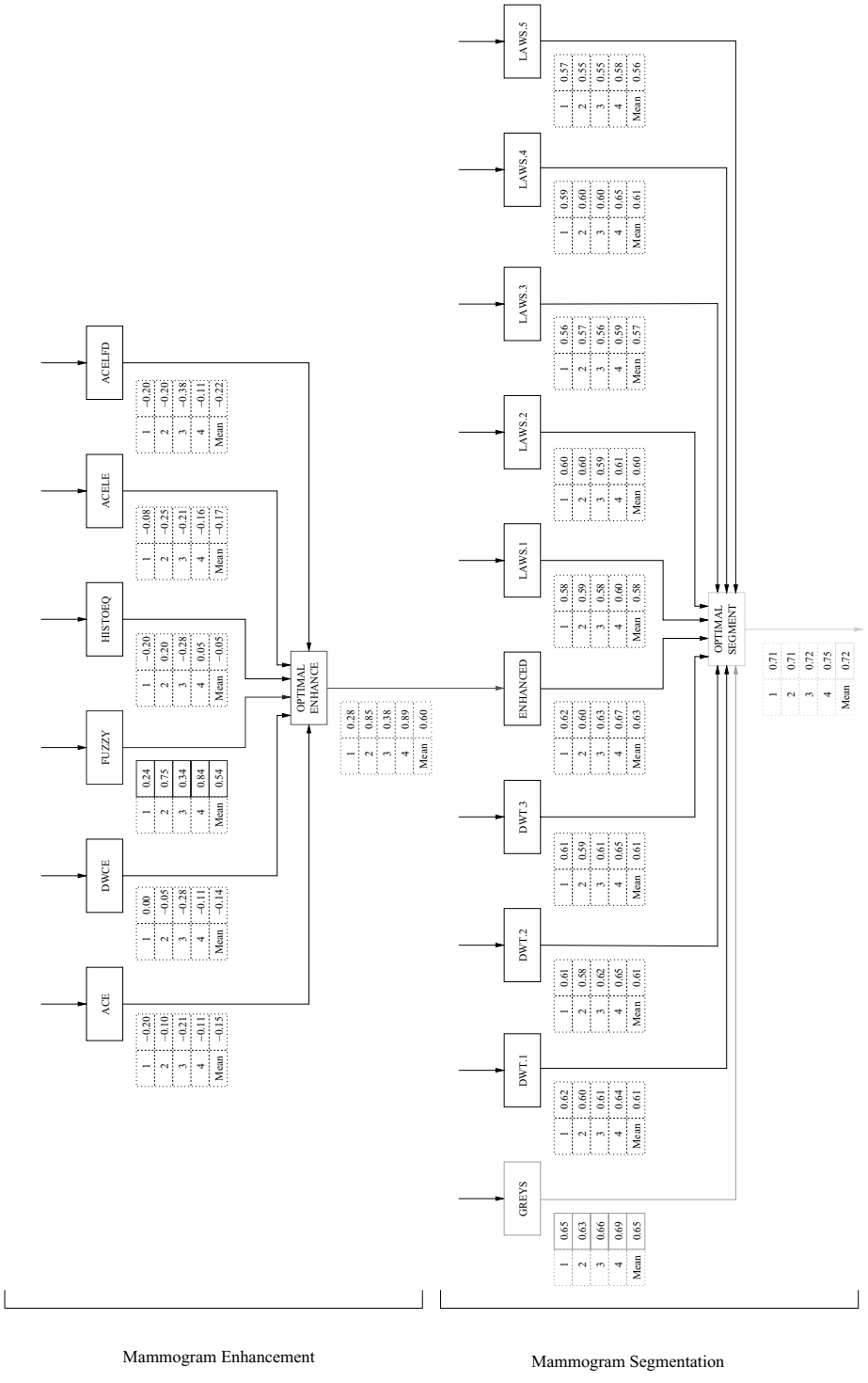


Figure 11.8: Evaluation of a given configuration of the adaptive knowledge-based model. Performance shown for each breast type for each component is interpreted as a percentage.

The following paragraphs briefly review the contrast enhancement and segmentation of digitized mammograms described in previous sections.

Contrast enhancement: The trained contrast enhancement knowledge-based component selects the optimal contrast enhancement method for a test mammogram, as one from a subset of six selected enhancement methods. Each of the enhancement methods has been described in section 11.3.2.1.1. The BPM strategy is used to implement the knowledge-based contrast enhancement component, and following training predicts the optimal enhancement method for a testing mammogram on the basis of an extracted feature vector. A different feature vector is used depending on the predicted breast type. A feature vector comprising a selected number of principal components $F_{\mathrm{BP}^{26}}$ is used for mammograms of breast types 1–3. For breast type 4, the complete feature vector $F_{\mathrm{BP}^{316}}$ is used.

Segmentation: To segment a mammogram, the semisupervised WGMM constrained with a $WGMM_S^{MRF}$ strategy is used. Ten different segmentation experts are trained and each one gives a segmentation decision for the test mammogram. The 10 experts have been trained to operate on specific groupings of input feature spaces. The experts for this configuration of the adaptive knowledge-based model are described in section 11.4.4. The decision of each expert is combined using a knowledge-based segmentation component implemented using the AWM described earlier. The AWM will predict the optimal blend of expert decisions to maximize the segmentation performance.

11.6.1.2 Results

This section presents the results from contrast enhancement and segmentation using 200 abnormal images, such that the image processing pipeline is constructed on the basis of the predicted breast type.

Knowledge-based contrast enhancement: From the results presented in Fig. 11.8, it can be seen the best performing expert is the FUZZY contrast enhancement method over all breast types. The average improvement in segmentation performance is 54% for all 200 abnormal images. Using the predicted optimal enhancement method from the knowledge-based contrast enhancement component, the average improvement in segmentation performance increases to 60%. The knowledge-based contrast enhancement component is determining the optimal enhancement based on component knowledge learnt during supervised

training. By utilizing this hidden knowledge, the resultant performance is improved compared with that obtained by simply using the single best contrast enhancement method, the FUZZY contrast enhancement method.

Knowledge-based mammogram segmentation: From the results presented in Fig. 11.8, it can be seen that the single best performing expert is the gray-scale contrast enhancement method for all predicted breast types. The mean A_Z value is 0.65 for all 200 abnormal images. Using predicted optimal enhancement method from the knowledge-based segmentation component, the mean A_Z value rises to 0.72. Clearly in combining the decisions of each expert, the knowledge-based segmentation component is performing better than that obtained by selecting the single best performing expert. The outcome of each segmentation expert is considered when forming the optimal segmentation. The statistically motivated AWM component, determines optimal weights for each segmentation expert that are most likely to have given rise to a resultant combined single segmentation. By doing this, the resulting performance is improved over all other segmentation experts.

11.6.1.3 Overlap Analysis of Segmentation Results

The results in the previous section show that the performance obtained following ROC analysis of the knowledge-based segmentation component is greater than that obtained from the best performing segmentation expert. By thresholding each probability image using a ROC operating point following optimal expert combination, region boundaries can be identified. In general, the ROC operating point [40] can be selected for each individual mammogram by associating a cost for a false positive, C_{FP}, and a false negative, C_{FN}. In this chapter, the operating point cannot be determined using this method. This is because the ground truth knowledge cannot be used during testing.

To determine an estimate of the operating point, the mean operating point is calculated from all mammograms contained within a training fold. Only mammograms that following segmentation, give lesion detection with an operating point greater than 0.95 are considered. The mean operating point is calculated from each training fold, for each breast type. To compute each operating point, the relative cost of a false positive is chosen as $C_{\mathrm{FP}} = 1$ and for a false negative $C_{\mathrm{FN}} = 20$. In addition, the probability of a positive outcome, $P(D+) = 0.03$, computed as the mean percentage of abnormal pixels in all training mammograms.

Table 11.16: Sensitivity and average number of false positives per image after segmentation of 200 abnormal images using adaptive knowledge-based model

Type	Sensitivity	FP/i
1	0.79 (0.57)	175.03 (163.31)
2	0.80 (0.58)	172.42 (132.09)
3	0.96 (0.72)	136.48 (132.55)
4	0.79 (0.66)	121.47 (158.79)
Mean	0.84 (0.63)	151.35 (146.73)

This is compared to the results (in brackets) from $WGMM_S^{MRF}$ experts combined using *AWM* model.

Following identification of region boundaries, overlap analysis is performed using the outcomes described in Table 11.2. The sensitivity results and average number of false-positive regions per image are shown in Table 11.16. The average number of false-positive regions per image decreases as the breast type increases. This can be attributed to the stricter ROC threshold used for thresholding these groups of probability images. Note that these results improve significantly on those without using adaptive model. This is because of the difficulty in selecting values for the costs C_{FP} and C_{FN} when setting the operating point. Different image segmentations result in different distributions of *a posteriori* estimates of positive (abnormal) and negative (normal) pixels. Ideally C_{FP} and C_{FN} need to be optimized on a per image basis, but this optimization is outside of the scope of this study.

11.6.2 Expert Contrast Enhancement and Segmentation of all Images with Adaptive Knowledge-Based Model

11.6.2.1 Dataset and Adaptive Knowledge-Based Model Configuration

In this section, the adaptive knowledge-based model is evaluated using the same configuration as described in the previous section and using exactly the same strategy for determining the segmentation operating point from an independent

Table 11.17: Frequency of normal and abnormal images

Type	Abnormal	Normal
1	53	54
2	20	20
3	28	36
4	99	90
Total	200	200

training set. The dataset is extended to include 200 normal images from four different breast types, 50 normal images drawn from each. The use of normal mammograms will demonstrate the specificity levels of the adaptive knowledge-based model. Table 11.17 shows the frequency of predicted breast groupings for normal and abnormal classes following breast type classification. The adaptive knowledge-based model is evaluated in its ability to provide an optimal segmentation for all the normal and abnormal mammograms.

11.6.2.2 Overlap Analysis of Segmentation Results

Using overlap analysis, both sensitivity and the average number of false-positives per image can be determined for each predicted breast group. The results from overlap analysis are shown in Table 11.18. From this table, it can be seen that the average number of false positives over all breast types has risen slightly with

Table 11.18: Sensitivity and average number of false positives per image after segmentation of 200 abnormal and 200 normal mammograms using the adaptive knowledge-based model

Type	Sensitivity	*FP/i*
1	0.79	207.26
2	0.80	162.68
3	0.96	161.86
4	0.80	136.45
Mean	0.84	167.01

the inclusion of the 200 normal mammograms compared with the results presented in Table 11.16. The aim of the false-positive reduction knowledge-based component described is to reduce the false-positive count, while maintaining sensitivity in the detection of lesions. The next section describes how this is achieved in this configuration of the adaptive knowledge-based model.

11.6.2.3 Reduction of False-Positives

False positives are initially reduced by removing regions with an area less than a predefined threshold T_{area}. We choose $T_{\text{area}} = 122$ pixels, thus any region less than 5 mm in diameter is removed. This approach is used here. From the remaining suspicious regions, features are extracted, and using a trained ANN classifier, a region is labelled as abnormal or normal.

Feature extraction: For those regions that remain following the application of the area test, the 316-dimensional feature vector described in [42] is extracted using the pixels comprising the region. To improve classifier generalization [32] unbiased PCA is used to map the 316-dimensional feature vector into a lower dimensional feature space. PCA is used on a per breast type basis, so that the number of principal components is selected independently for each breast type. Using this approach, for each predicted breast type, the number of principal components selected are as follows: (type 1, 37 components; type 2, 33 components; type 3, 35 components; type 4, 41 components). From this table it can be seen that the highest dimensional feature space results from the densest breast types (type 4), which are the generally the hardest to interpret by an expert radiologist [37].

Model order selection: In order to maximize the performance of each ANN for each predicted breast type, model order selection of the ANN classifier is performed. By varying the number of hidden nodes and performing a classification on all suspicious region, ROC can be performed and the area under the ROC curve (A_Z) computed. The optimal number of hidden nodes is determined as that maximising the A_Z value.

11.6.2.4 Results

This section presents the results from applying the false-positive reduction methodology on the suspicious regions resulting from the knowledge-based

Table 11.19: Sensitivity and average number of false-positives per image for 200 abnormal and 200 normal images

	After segmentation		After segmentation prefiltering		After falsepositive reduction (OP)	
Breast type	Sensitivity	*FP/i*	Sensitivity	*FP/i*	Sensitivity	*FP/i*
1	0.79	207.26	0.77	8.98	0.64	3.41
2	0.80	162.48	0.80	10.95	0.70	7.05
3	0.96	161.86	0.89	7.96	0.89	7.76
4	0.80	136.45	0.76	6.71	0.76	7.63
Mean	0.84	167.01	0.81	8.65	0.75	6.46

Values after segmentation, after region prefiltering, and after false-positive reduction using optimized classifier at ROC operating point, each by breast type (FP/i = average number of false-positive per image, OP = operating point).

contrast enhancement and segmentation components. Table 11.19 shows the sensitivity in the detection of breast lesions and average number of false-positive regions over all mammograms of each predicted breast type. The first column of shows performance after segmentation, the second column shows the results after region prefiltering, and the final column shows the results following classification using the optimized ANN.

The aim of false-positive reduction is to reduce the average number of false-positive regions per image while maintaining sensitivity levels. The prefiltering stage can be seen from the results as being very effective in reducing the false-positive count. After region prefiltering sensitivity has dropped by just over 3.5% and the average number of false-positive regions to 8.65 for all 400 images. The mean sensitivity drops more sharply when reducing false-positive regions with the trained ANNs. The performance of each ANN is reported at the operating point on the ROC curve where the cost of a false positive and false negative are set equal ($C_{\mathrm{FP}} = C_{\mathrm{FN}} = 1$) and the *a priori* probability of a positive case is set, $P(D+) = 0.5$. These costs and priors can be adjusted by the expert radiologist to reflect the required level of sensitivity. Optimization of their values is outside the scope of this study. The largest reduction in false-positive regions using the ANN is seen for the fatty, type 1 breasts. For the denser breasts, types 3 and 4, the operating point selected maintains the level of sensitivity but does not significantly reduce the false-positive count. In fact for the densest breasts, the false-positive count rises indicating the nontrivial nature of this classification problem.

11.6.3 Key Observations

This section has presented a configuration of the adaptive knowledge-based model. The performance of the model has been evaluated on a dataset of 200 abnormal mammograms from four different breast types. The aim of the evaluation has been to demonstrate the utility of the model compared with that of individual experts. Following this, the performance of a specific configuration of the adaptive knowledge-based model has been evaluated on a dataset of 400 mammograms, comprising 200 abnormal and 200 normal images. From these evaluations, the following key observations can be made:

1. *Utility of knowledge-based contrast enhancement component*: Using a dataset of 200 abnormal mammograms, the utility of the knowledge-based contrast enhancement expert has been demonstrated to be greater than that of the best performing expert contrast enhancement method. Using the predicted optimal contrast enhancement method in image segmentation results in a 60% improvement in the detection of abnormal regions over the original segmentation. This is compared to a 54% improvement from the single best performing expert, the FUZZY contrast enhancement method.

2. *Utility of knowledge-based segmentation component*: By optimally combining the segmentation outcomes of 10 different segmentation experts, each operating on a unique feature space partition, the knowledge based segmentation component resulted in a mean ROC A_Z value of 0.72 for 200 mammograms from four breast types. This is compared to the best performing gray-scale segmentation expert reporting a mean A_Z value of 0.65.

3. *Utility of adaptive knowledge-based model in presence of normal mammograms*: Evaluation of the performance of this configuration of the adaptive knowledge-based model on a dataset of 400 mammograms comprising 200 abnormal and 200 normal images results in a segmentation sensitivity of 0.84 for the detection of breast lesion with 167.01 false-positive regions per image. This demonstrates a high level of sensitivity in the presence of a complete spectrum of mammogram types.

4. *False-positive reduction*: The results following region prefiltering in the false-positive reduction methodology demonstrate the utility of the region

size thresholding strategy. Following classification by each trained optimized ANN results in a sensitivity of 0.75 with 6.46 false-positive regions per image.

11.7 Conclusions

In this chapter we have presented a framework for adaptive selection of image processing components based on image properties. Throughout the study, we have evaluated the different components and the overall model on the same dataset in order to produce a consistent and comparable set of results. The framework presented here has generic applicability to medical imaging applications and we are confident that further research will involve such knowledge-based approaches.

Questions

1. *What is the essence of the chapter?*
2. *What are the two main areas used in this chapter when it comes to X-ray breast imaging?*
3. *What are the different measures used for X-ray breast "contrast enhancement"? Discuss each of them.*
4. *Show how the knowledge-based system works for the contrast enhancement.*
5. *What is the role of image segmentation here and how it is done?*
6. *What is double network mapping (DNM)?*
7. *What is breast profile mapping (BPM)? Discuss in detail.*
8. *Discuss the weighted GMM/MRF model of segmentation of breast masses in X-ray images. State mathematically and then discuss the pseudo algorithm.*
9. *Compare the four models* $\left(WGMM_S, WGMM_U, WGMM_S^{MRF}, WGMM_U^{MRF}\right)$.
10. *List some key observations in adaptive knowledge-based model.*

Bibliography

[1] Rangayyan, R. M. *et al.*, Improvement of sensitivity of breast cancer diagnosis with adaptive neighbourhood contrast enhancement of mammograms, IEEE Trans. Inf. Tech. Biomed., Vol. 1, No. 3, pp. 161–169, 1997.

[2] Petrick, N. *et al.*, Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification, Med. Phys., Vol. 23, pp. 1685–1696, 1996.

[3] Li, L., Qian, W., and Clarke, L. P., Digital mammography: Computer assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms, Acad. Radiol., Vol. 4, No. 11, pp. 724–731, 1997.

[4] Sahiner, B. *et al.*, Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue, Med. Phys., Vol. 23, No. 10, pp. 1671–1683, 1996.

[5] Sahiner, B. *et al.*, Computer-aided characterisation of mammographic masses: Accuracy of mass segmentation and its effects on characterisation, IEEE Trans. Med. Imaging, Vol. 20, No. 12, pp. 1275–1284, 2001.

[6] Polakowski, W. E. *et al.*, Computer aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency, IEEE Trans. Med. Imaging, Vol. 16, pp. 811–819, 1997.

[7] Yin, F. F. *et al.*, Comparison of bilateral subtraction and single image processing techniques in the computerised detection of mammographic masses, Invest. Radiol., Vol. 28, No. 6, pp. 473–481, 1993.

[8] Singh, S. and Al-Mansoori, R., Identification of region of interest in digital mammograms, J. Intell. Syst., Vol. 10, No. 2, pp. 183–210, 2000.

[9] Zheng, B., Chang, Y., and Gur, D., Adaptive computer-aided diagnosis scheme of digitised mammograms, Acad. Radiol., vol. 3, pp. 806–814, 1996.

[10] Matsubara, T. *et al.*, Development of new schemes for detection and analysis of mammographic masses, In: Proceedings of the International Conference on Intelligent Information Systems, 1997.

[11] Lai, S. and Fang, M., Adaptive medical image visualisation based on hierarchical neural networks and intelligent decision fusion, In: Proceedings of IEEE Signal Processing Society Workshop, pp. 438–447, 1998.

[12] Lai, S. and Fang, M., A hierarchical neural network algorithm for robust and automatic windowing of MR images, Artif. Intell. Med., Vol. 19, pp. 97–119, 2000.

[13] Pitiot, A., Toga, A. W., Ayache, N., and Thompson, P., Texture based MRI segmentation with a two stage hybrid neural classifier, In: Proceedings of IEEE IJCNN Conference, 2002, Vol. 3, pp. 2053–2058.

[14] Sha, D. D. and Sutton, J. P., Towards automated enhancement, segmentation and classification of digital brain images using networks of networks, Inf. Sci., Vol. 138, pp. 45–77, 2001.

[15] Fenster, S. D. and Kender, J. R., Sectored snakes: Evaluating learned-energy segmentations, IEEE Trans. Med. Imaging, Vol. 23, No. 9, pp. 1028–2034, 2001.

[16] Perner, P., An architecture for a CBR image segmentation system, Eng. Appl. Artif. Intell., Vol. 12, No. 6, pp. 749–759, 1999.

[17] Guan, L., Anderson, J. A., and Sutton, J. P., A network of networks processing model for image regularisation, IEEE Trans. Neural Networks, Vol. 8, No. 1, pp. 169–174, 1997.

[18] Bovis, K. J., An Adaptive Knowledge-Based Model for Detecting Masses in Screening Mammograms, Ph.D. Thesis, Department of Computer Science, University of Exeter, 2003.

[19] Singh, S. and Bovis, K. J., Digital mammography segmentation, In: Advanced Algorithmic Approach to Medical Image Segmentation: State-of-the-Art Application in Cardiology, Neurology, Mammography and Pathology, Suri, J., Setarehdan, S. K., and Singh, S., eds., Springer-Verlag, Berlin, pp. 440–540, 2001.

[20] Bovis, K. J. and Singh, S., Enhancement technique evaluation using quantitative measures on digital mammograms, In: Proceedings of the 5th International Workshop on Digital Mammography, Toronto, Canada, Yaffe, M. J., ed., Medical Physics Publishing, pp. 547–553, 2000.

[21] Weszka, J. S. and Rosenfeld, A., A comparison study of texture measures for terrain classification, In: Proceedings of Conference in Computer Graphics, 1975, pp. 62–64.

[22] Hair, J., Anderson, R., and Tatham, R., Multivariate data analysis, 1998.

[23] Webb, A., Statistical Pattern Recognition, Arnold, 1999.

[24] Zhang, Y., Brady, M., and Smith, S., Segmentation of brain MR images through a hidden Markov random field model and the expectation minimisation algorithm, IEEE Trans. Med. Imaging, Vol. 20, No. 1, pp. 45–57, 2001.

[25] Kallergi, M., Carney, G. M., and Gaviria, J., Evaluating the performance of detection algorithms in digital mammography, Med. Phys., Vol. 26, No. 2, pp. 267–275, 1999.

[26] Bovis, K. J. and Singh, S., Classification of mammographic breast density using a combined classifier paradigm, In: Medical Image Understanding and Analysis (MIUA) Conference, Portsmouth, July 22–23, 2002.

[27] Duda, R. O., Hart, P. E., and Stork, D. G., Pattern Classification, Wiley, Neq York, 2001.

[28] Sonka, M., Hlavac, V., and Boyle, R., Image Processing, Analysis and Machine Vision, PSW Publishing, 1999.

[29] Besag, J., On the statistical analysis of dirty pictures, J. Roy. Soc. B, Vol. 48, No. 3, pp. 259–302, 1986.

[30] Dubes, R. C. and Jain, A. K., Random field models in image analysis, J. Appl. Stat., Vol. 16, No. 2, pp. 131–163, 1989.

[31] Geman, S. and Geman, D., Stochastic relaxation Gibbs distributions and Bayesian restoration of image, IEEE Trans. PAMI, Vol. 6, No.6, pp. 721–741, 1984.

[32] Bishop, C. M., Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[33] Kittler, J., Combining classifiers: A theoretical framework, Patt. Anal. Appl., Vol. 1, No. 1, pp. 18–27, 1998.

[34] Jacobs, R. A. *et. al.*, Adaptive mixtures of local experts, Neural Comput., Vol. 3, pp. 79–87, 1991

[35] Dempster, A. P., Laird, N. M., and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc. B, Vol. 39, pp. 1–38, 1977.

[36] Yin, F. F. *et al.* Computerised detection of masses in digital mammograms: Analysis of bilateral subtraction images, Med. Phys., Vol. 18, No. 5, pp. 955–963, 1991.

[37] Kopans, D., Breast Imaging, Lippincott-Raven, 1998.

[38] Haralick *et al.*, 1973.

[39] Wei *et al.*, 1997.

[40] Metz, C. E., Basic principles of ROC analysis, Sem. Nuclear Med., Vol. 8, No. 4, pp. 283–298, 1978.

[41] Kupinski, M. A. and Anastasio, M. A., Multi-objective genetic optimisation of diagnostic classifiers with implications for generating receiver operating characteristic curves, IEEE Trans. Med. Imaging, Vol. 18, No. 8, pp. 675–685, 1999.

[42] Bovis, K. J. and Singh, S., Learning the optimal contrast enhancement of mammographic breast masses, In: Proceedings of the 6th International Workshop on Digital Mammography, Bremen, Germany, June, 22–25, 2002, Springer, Berlin, pp. 179–181, 2002.

Chapter 12

Simultaneous Fuzzy Segmentation of Medical Images

Gabor T. Herman[1] *and Bruno M. Carvalho*[2]

12.1 Introduction

Digital image segmentation is the process of assigning distinct labels to different objects in an image. The level of detail indicated by the labeling is related to the application at hand. To perform object identification in digital or continuous, moving or still images, humans make use of high-level reasoning and knowledge, as well as of different visual cues, such as shadowing, occlusion, parallax motion, and the relative size of objects. Aside from the difficulty of inserting this type of reasoning into a computer program, the task of segmenting out an object from its background in an image becomes particularly hard for a computer when, instead of the brightness values, what distinguishes the object from the background is some textural property, or when the image is corrupted by noise and/or inhomogeneos illumination.

Segmentation algorithms can be divided into three categories according to their user-program interactivity: manual, semiautomatic, and automatic. In manual algorithms, users can make use of some computer routines (e.g., drawing tools) to isolate and segment one or more objects. In semiautomatic algorithms, users usually select some points or areas that will be used to collect information for the characterization of the objects to be segmented. Finally, the automatic

[1]Doctoral Program in Computer Science, The Graduate Center, CUNY, 365 5th Avenue, New York, NY

[2]Department of Computer and Information Science, University of Pennsylvania, Levine Hall, 3330 Walnut Street, Philadelphia, PA

algorithms perform the whole segmentation process without any user intervention, usually by obtaining all information necessary to perform the segmentation from prior knowledge about the class of problems to which the segmentation at hand belongs.

Algorithms can be also be classified according to how they solve the segmentation problem. Point-based algorithms make a local decision about a point's membership to an object. This decision can be based solely on the point's brightness value or on the brightness values of a small neighborhood surrounding the point. A widely used and very simple point-based segmentation algorithm is thresholding, where a user selects one or two brightness values that are interpreted as lower and/or upper values of the brightness of the object to be segmented. Then, all pixels whose values are in the specified brightness range are considered to be part of the object. It is easy to see that algorithms of this type are very sensitive to noise, to inhomogeneous illumination, and are not appropriate for segmenting textured objects.

Edge-based segmentation algorithms usually work in two steps by first detecting edges in the image and then grouping or linking them into boundaries of objects based on the orientation of the edges and on prior knowledge regarding the expected shape of objects. Common edge detection procedures include the use of gradient operators, Laplacians or the Canny edge detector [1], while edge linking can be performed locally by searching small local pixel neighborhoods or globally by making use of the Hough Transform [2], for example. Other edge-based segmentation algorithms use active contour models, such as snakes [3] or balloons [4]. Snakes are energy-minimizing splines guided by external constraint forces and pushed by image forces (edges) toward image features, while balloons use image forces to stop their inflated curve models on image features. There are also global optimization algorithms [5–7] that segment images by minimizing various energy functions defined in terms of pixel labels and prior knowledge.

Region-based algorithms are subdivided into region growing and split-and-merge algorithms. Region growing algorithms, as the name suggests, start with preselected *seed* points forming the initial regions that grow according to some predefined rules until the whole image is labeled. Split-and-merge algorithms begin by subdividing an image into arbitrary disjoint regions, and then split and/or merge them repeatedly until some preset conditions are satisfied. The methods of balloons [4, 8] and level sets [9, 10] can also be considered region

growing methods since they make use of contour models that inflate from an initial position to segment objects in a scene.

In this work we present a multiseeded fuzzy segmentation algorithm, which is a greedy semiautomatic region growing algorithm based on the fuzzy segmentation algorithm of [11] but is capable of efficiently segmenting multiple objects simultaneously.

12.2 Fuzzy Segmentation

If what distinguishes objects in an image are not the exact values assigned to the pixels but rather some textural property (as it is the case for images containing random noise and/or shading), then fuzzy connectedness can be usefully employed to achieve segmentation (see [12–17] and their references). Fuzzy connectedness was explicitly introduced by Rosenfeld [18], but it had been foreshadowed earlier (for example by the "Minimum Method" in [13]). Our approach is based on that advocated in [11], but is generalized to arbitrary digital spaces [19].

A *digital space* is a pair (V, π), where V is a set and π is a symmetric binary relation on V such that V is connected under π. A *picture* over this digital space is a triple (V, π, f), where f maps V into the real numbers. Because of the nature of the applications that we have in mind, we refer to elements of V as *spels*, which is short for *spatial elements* [19]. In this paper we assume that π is antireflexive (i.e., that, for all $c \in V$, $(c, c) \notin \pi$) and we use $N(c)$ to denote the *neighborhood* of c that consists of c itself and all $d \in V$, such that $(c, d) \in \pi$. If $(c, d) \in \pi$, we say that c and d are *adjacent*. The spels can be pixels of an image (as in [11, 12, 14, 16–18, 20]), but they can also be dots in the plane (as in [21, 22]), or any variety of other things. The theory and algorithm presented here will be independent of the specifics of the application area. They are in particular applicable to data clustering [23] in general, and so their range of usefulness goes far beyond just image segmentation and includes such distant areas of endeavor as psychology [13] and statistics [24].

The basic concept that we are generalizing here is that of *fuzzy connectedness*: to every ordered pair (c, d) of spels, it assigns a real number not less than 0 and not greater than 1. This indeed is an example of a fuzzy set (as it is normally defined in the literature [25]): the fuzzy set in question is "the set of

connected pairs" and the *grade of membership* of (c, d) in this set is the fuzzy connectedness of c to d. In the approach used below, fuzzy connectedness is defined in the following general manner.

We call a sequence of distinct spels a *chain*; its *links* are the ordered pairs of consecutive spels in the sequence. We define the *ψ-strength* of a link to be the appropriate value of a *fuzzy spel affinity* function $\psi : V^2 \to [0, 1]$, i.e., a function that assigns a value between 0 and 1 to every pair of spels in V. For example, if the set of spels V is a finite set of dots in the plane, we may define the strength of the link from one dot to another as the reciprocal of the distance between them (we need to make the unit of distance such that all distinct dots are at least one unit from each other). A chain is formed by one or more links and the *ψ-strength of a chain* is the ψ-strength of its weakest link; the ψ-strength of a chain with only one spel in it is 1 by definition. A set $U(\subseteq V)$ is said to be *ψ-connected* if, for every pair of spels in U, there is a chain in U of positive ψ-strength from the first to the second spel of the pair. As we will see later, for the purpose of fuzzy segmentation of images, the strength of any link of one pixel to another can often be automatically defined based on statistical properties of the links within regions identified by the user as belonging to the object of interest.

We associate with the fuzzy spel affinity function ψ a *fuzzy connectedness* function $\mu_\psi : V^2 \to [0, 1]$ defined by

$$\mu_\psi(c, d) = \max_{\substack{<c^{(0)},\ldots,c^{(K)}> \in V^{K+1} \\ c^{(0)}=c,\, c^{(K)}=d \text{ and } c^i \neq c^j \text{, if } i \neq j}} \min_{1 \leq k \leq K} \psi(c^{(k-1)}, c^{(k)}), \tag{12.1}$$

i.e., the ψ-strength of the strongest chain from c to d. We then define the *ψ-connectedness map* f of a set V for a seed spel o by the fuzzy connectedness values of o to c ($f(c) = \mu_\psi(o, c)$), for all $c \in V$. A *hard object* C is then defined based on the ψ-connectedness map by selecting a threshold t and associating with C all spels c for which $f(c)$ is above the threshold, i.e., $C = \{c \mid c \in V,\ f(c) \geq t\}$.

The algorithm proposed in [11] for obtaining a fuzzy connectedness map uses the concept of dynamic programming and has the characteristic that a single spel can be put into a spel queue O (that holds the spels waiting to be considered in the search for optimal chains) many times. This seemed to us an unnecessary inefficiency. In [12] we investigated the use of so-called *greedy algorithms* [26] for computing the fuzzy connectedness map. We observed that if we treat the set V as a connected graph and we consider the *cost* of the arc (c, d) to be $1 - \psi(c, d)$, some of the graph algorithms for finding *shortest paths* could be applied to this

problem. We showed that both Dijkstra's and Prim's algorithms can be used for computing the fuzzy connectedness map of an image faster than the previously used dynamic programming algorithm. In the experiments reported in [12] we achieved an average speedup of 8.2 times (over the algorithm of [11]) when using Dijkstra's or Prim's Algorithms for computing the connectedness maps for a set of images with the same size as the image shown in Fig. 12.1 ($|V| = 10{,}621$).

To obtain a version of Dijkstra's algorithm for computing the fuzzy connectedness map we only have to make two changes to the algorithm of [11]. First, we make O a set instead of a queue, and second, when we remove a spel from O, we remove the spel d for which $f(d)$ is maximal (greedy step). If a spel c is already in O it is not reinserted since O is now a set. This is the reason why this greedy algorithm is more efficient than the dynamic programming algorithm, in which a spel may be inserted into O many times. In order to implement efficiently the removal of d with the maximal $f(d)$, we make use of a priority queue, in our case a binary heap, that maintains a partial ordering of the elements in O [26].

In order to apply the algorithms mentioned above to image segmentation, we have to define the fuzzy spel affinity ψ. Usually this is done by a computer program, based on some minimal information supplied by a user [11, 12, 19]. The underlying idea is that, even though the user most likely will not be able to define mathematically the characteristics of the object of interest, it is quite easy for him/her to select a spel belonging to it. The program will then compute some statistics based on the neighborhood of the selected spel and use these statistics to compute the fuzzy spel affinity ψ. We now make a sample methodology we have been using to achieve this.

For a picture (V, π, I) and selected spel o, we define ψ by

$$\psi(c,d) = \begin{cases} \frac{[g_1(I(c)+I(d))+g_2(|I(c)-I(d)|)]}{2} & \text{if } (c,d) \in \pi, \\ 0 & \text{otherwise,} \end{cases} \tag{12.2}$$

where, for $i \in \{1, 2\}$,

$$g_i(x) = e^{-\frac{(x-m_i)^2}{2\sigma_i^2}}. \tag{12.3}$$

The values for m_i and σ_i are computed using the spels in the neighborhood of o: m_1 and σ_1 are defined as the mean and standard deviation, respectively, of $I(c) + I(d)$ over all adjacent spels c and d in $N(o)$ and m_2 and σ_2 are defined to be the mean and standard deviation, respectively, of $|I(c) - I(d)|$ over all adjacent

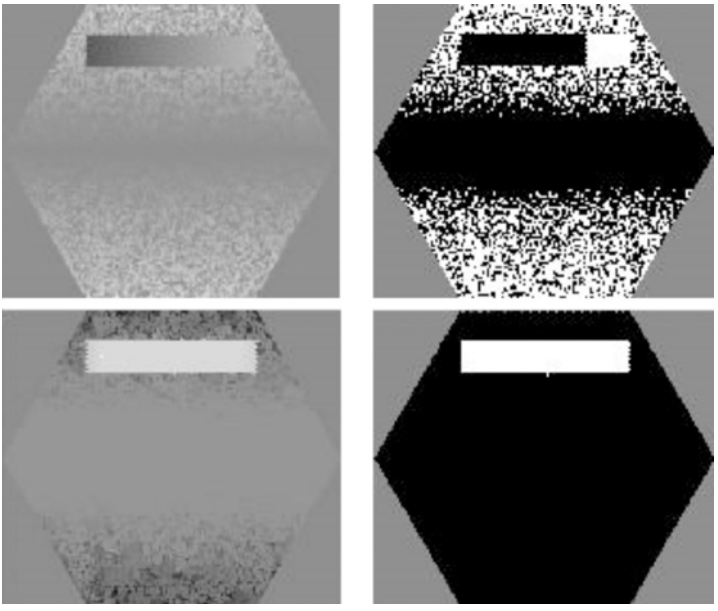


Figure 12.1: A mathematically defined image (top-left) on a hexagonal grid was segmented using thresholding (top-right) and fuzzy segmentation (bottom row).

spels c and d in $N(o)$. This means that for any pair (c, d) of adjacent spels, their fuzzy spel affinity will be large if both $I(c) + I(d)$ and $|I(c) - I(d)|$ have values similar to those in the neighborhood of the selected spel. This definition reflects the fact that in many applications both the values assigned by I to spels and the differences between the values assigned to neighboring spels are important for distinguishing objects in an image.

As an example, the top-left image in Fig. 12.1 was mathematically defined on the hexagonal grid (each element of V is a hexagon and all of them are arranged on an enclosing hexagon with $|V| = 10{,}621$), and the user was asked to select a spel that is located inside the object to be segmented. The object in question is the rectangular region in the upper half of the image with slowly increasing brightness from left to right. In this example, two hexagons are considered to be adjacent if, and only if, they share an edge (thus the neighborhood of any interior hexagon consists of seven hexagons), and $I(c)$ is the gray value assigned to the hexagon c.

The image on the top-right of Fig. 12.1 is the result of thresholding the original image at some level. Note that because of the brightness variation inside the

object that we wish to segment (the horizontal stripe near the top of the image) there is no threshold level that can successfully segment it from the background. When using the fuzzy segmentation algorithm, the user chose a point belonging to the object (the brightest point in the lower-left image) that is used to identify the neighborhood over which information is collected regarding the characteristics of the object, to be used in Eqs. (12.2) and (12.3). The resulting fuzzy spel affinity ψ is then used to produce the connectedness map f shown in the lower-left image (note that (V, π, f) is also a picture over the digital space (V, π)), which is then thresholded to produce the successful final segmentation shown in the lower-right image (the hexagons belonging to the resulting hard object are shown white).

12.3 Multiseeded Fuzzy Segmentation

The idea behind multiseeded fuzzy segmentation is to generalize the approach described in the previous section to multiple objects: each of the objects in the image has its own definition of strength for the links and its own set of seed spels. Each of the objects is then defined as the collection of those spels that are connected entirely within the object to one of its own seed spels in a stronger way than to any of the other seed spels. This intuitive notion will be made precise. An essential feature of our approach is that it does *not* simply calculate, for every spel, the grade of membership to each of the individual objects of that spel and then assigns the spel to the object for which its grade of membership is maximal (such an algorithm is discussed in [27]). The reason for this is that if a spel is separated from the seed points of Object 1 by spels belonging to Object 2, then it should not be assigned to Object 1. The gestalt that we are trying to capture here is a segmentation in which the chains that determine "belonging to an object" must lie entirely in that object.

A potentially time-consuming step in finding such objects is the calculation of the multiple fuzzy connectedness of all the spels to the seed spels. We devised a greedy and efficient algorithm that provides the desired segmentation. We demonstrate its performance on various mathematically-defined and physically obtained (real) images. The output of the process is a segmentation into fuzzy sets in the classical sense ([25], p. 39) that, for each spel, we also produce a "grade of membership" in the object(s) to which it belongs.

Similarly to the method presented in last section, we rely on the user of our method to identify seed spels that definitely belong to the various objects into which we desire to segment the images, and we suggest (as other advocates of segmentation based on fuzzy connectedness have done before us) that the user-selected seed spels can be used for automatic calculation of the definitions of the strengths of links in each one of the objects. Since our choice implies that the output of our algorithm is user-dependent, we report on experiments (in which five users segmented five images, each five times) that validate the accuracy and robustness of our approach.

12.3.1 Theory

For a positive integer M, an *M-semisegmentation* of a set V of spels is a function σ that maps each $c \in V$ into an $(M+1)$-dimensional vector $\sigma^c = (\sigma_0^c, \sigma_1^c, \ldots, \sigma_M^c)$, such that

1. $\sigma_0^c \in [0, 1]$ (i.e., it is nonnegative but not greater than 1),
2. for each m the value of σ_m^c is either 0 or σ_0^c, and
3. for at least one m in the range $1 \leq m \leq M$, $\sigma_m^c = \sigma_0^c$.

We say that σ is an *M-segmentation* if, for every spel c, σ_0^c is positive.

If there are multiple objects to be segmented, it is reasonable that each should have its own fuzzy spel affinity. For images this idea is discussed in [27], here it is generalized by the following concepts. An *M-fuzzy graph* is a pair (V, Ψ), where V is a nonempty finite set and $\Psi = (\psi_1, \ldots, \psi_M)$ with ψ_m (for $1 \leq m \leq M$) being a fuzzy spel affinity. A *seeded* M-fuzzy graph is a triple $(V, \Psi, \mathcal{V})$, where (V, Ψ) is an M-fuzzy graph and $\mathcal{V} = (V_1, \ldots, V_M)$, where $V_m \subseteq V$, for $1 \leq m \leq M$. A seeded M-fuzzy graph is *connectable* if

1. the set V is $(\min_{1 \leq m \leq M} \psi_m)$-connected (we define $\left(\min_{1 \leq m \leq M} \psi_m\right)(c, d) = \min_{1 \leq m \leq M} \psi_m(c, d)$) and
2. $V_m \neq \emptyset$, for at least one m, $1 \leq m \leq M$.

For an M-semisegmentation σ of V and for $1 \leq m \leq M$, the chain $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ is said to be a *σm-chain* if $\sigma_m^{c^{(k)}} > 0$, for $0 \leq k \leq K$. Furthermore, for $W \subseteq V$ and $c \in V$, we use $\mu_{\sigma,m,W}(c)$ to denote the maximal ψ_m-strength of a σm-chain from a spel in W to c. (This is 0 if there is no such chain.)

Theorem 1.1. *If* $(V, \Psi, \mathcal{V})$ *is a seeded M-fuzzy graph, then*

(i) there exists an M-semisegmentation σ *of* V *with the following property: for every* $c \in V$, *if for* $1 \le n \le M$

$$s_n^c = \begin{cases} 1 & \text{if } c \in V_n, \\ max_{d \in V}(min(\mu_{\sigma,n,V_n}(d), \psi_n(d,c))) & \text{otherwise}, \end{cases} \tag{12.4}$$

then for $1 \le m \le M$

$$\sigma_m^c = \begin{cases} s_m^c, & \text{if } s_m^c \ge s_n^c \quad \text{for } 1 \le n \le M, \\ 0 & \text{otherwise}, \end{cases} \tag{12.5}$$

(ii) this M-semisegmentation is unique, and

(iii) it is an M-segmentation, provided that $(V, \Psi, \mathcal{V})$ *is connectable.*

Before discussing the validity of Theorem 1.1, we discuss in less mathematical terms what it says. The property stated in Theorem 1.1 is a reasonable one, as can be seen in Fig. 12.2. Let c be an arbitrary spel and suppose that σ^d is known for all other spels d. Then, for $1 \le n \le M$ ($M = 3$ in Fig. 12.2), the s_n^c of Eq. (12.4) is the maximal ψ_n-strength of a chain $\langle d^{(0)}, \ldots, d^{(L)}, c \rangle$ from a (seed) spel in V_n to c such that $\sigma_n^{d^{(l)}} > 0$ (i.e., $d^{(l)}$ belongs to the nth object), for $0 \le l \le L$. (s_n^c is defined to be 0 if there is no such chain.) Intuitively, the mth object can

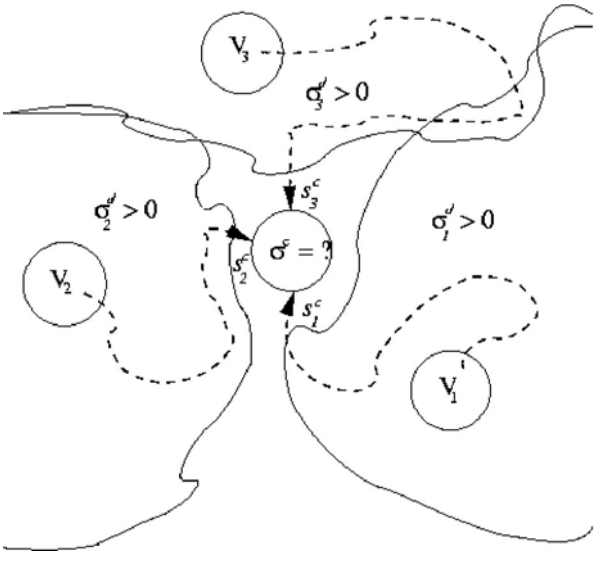


Figure 12.2: Illustration of the desirability of the M-segmentation whose existence (and uniqueness) is guaranteed by Theorem 1.1.

"claim" that c belongs to it if, and only if, s_m^c is maximal and is greater than 0. This is indeed how things get sorted out in Eq. (12.5): σ_m^c has a positive value only for such objects. Furthermore, this is a localized property in the following sense: for a fixed spel c we can work out the values of the s_n^c using Eq. (12.4) and what we request is that, at that spel c, Eq. (12.5) be satisfied. What Theorem 1.1 says that there is one, and only one, M-semisegmentation that satisfies this property simultaneously everywhere, and that this M-semisegmentation is in fact an M-segmentation, provided that the seeded M-fuzzy graph is connectable.

Now we illustrate the Theorem 1.1 for the seeded two-fuzzy graph $\left(\overline{V}, \overline{\Psi}, \overline{\mathcal{V}}\right)$ defined by $\overline{V} = \{(-1), (0), (1)\}$ and $\overline{\Psi} = \left(\overline{\psi}_1, \overline{\psi}_2\right)$, where $\overline{\psi}_1$ and $\overline{\psi}_2$ are the reflexive and symmetric fuzzy spel affinity functions (i.e., $\overline{\psi}_m(c, c) = 1$ and $\overline{\psi}_m(c, d) = \overline{\psi}_m(d, c)$ for $1 \leq m \leq 2$ and $c, d \in \overline{V}$) defined by the additional conditions $\overline{\psi}_1((-1), (0)) = 0.5$, $\overline{\psi}_1((0), (1)) = 0.25$, $\overline{\psi}_1((-1), (1)) = 0$, $\overline{\psi}_2((-1), (0)) = 0.5$, $\overline{\psi}_2((0), (1)) = 0.5$, $\overline{\psi}_2((-1), (1)) = 0$, and $\overline{\mathcal{V}} = (\{(0)\}, \{(-1)\})$. The two-segmentation $\overline{\sigma}$ of $\overline{V}$ that satisfies Theorem 1.1 is given by $\overline{\sigma}^{(-1)} = (1, 0, 1)$, $\overline{\sigma}^{(0)} = (1, 1, 0)$, and $\overline{\sigma}^{(1)} = (0.25, 0.25, 0)$. To test this suppose, for example, that we have been informed that $\overline{\sigma}^{(-1)} = (1, 0, 1)$ and $\overline{\sigma}^{(0)} = (1, 1, 0)$ and we wish to use the Theorem 1.1 to determine $\overline{\sigma}^{(1)}$. We find that $s_1^{(1)} = 0.25$ (obtained by the choice $d = (0)$) and $s_2^{(1)} = 0$ (if we choose in Eq. (12.4) d to be (-1), then $\overline{\psi}_2((-1), (1)) = 0$, if we choose it to be (0), then $\mu_{\overline{\sigma},2,(-1)}(0) = 0$ since there is no $\overline{\sigma}2$-chain containing (0), because $\overline{\sigma}_2^{(0)} = 0$). Hence Eq. (12.5) tells us that indeed $\overline{\sigma}^{(1)} = (0.25, 0.25, 0)$. Note that there is a chain $\langle(-1), (0), (1)\rangle$ of $\overline{\psi}_2$-strength 0.5 from the seed spel (-1) of Object 2 to (1), while the maximal $\overline{\psi}_1$-strength of any chain from the seed spel (0) of Object 1 to (1) is only 0.25; nevertheless, (1) is assigned to Object 1 by Theorem 1.1, since the fact that (0) is a seed spel of Object 1 prevents it (for the given $\overline{\Psi}$) from being also in Object 2, and so the chain $\langle(-1), (0), (1)\rangle$ is "blocked" from being a $\overline{\sigma}2$-chain.

An intuitive picture of the inductive definition of the M-semisegmentation of Theorem 1.1 is given by the following description of a military exercise. There are M armies (one corresponding to each object) competing for the control of N castles that are connected by one-way roads between them. Initially all armies have full strength (defined to be 1) and they occupy their respective castles (seed spels). All armies try to increase their respective territories, but the moving from a castle to another one reduces the strength of the soldiers to the minimum of their strength at the previous castle and the affinity (for that army or object) of

the road connecting the castles. The affinities of the roads for the various armies are fixed for the duration of the exercise.

All through the exercise each castle will also have have a strength assigned to it, which is a real value in the range [0, 1]. The strength of a castle may increase as the exercise proceeds. Also, at any time, each castle may be occupied by one or more of the armies.

The objective of the exercise is to see how the final configuration of occupied castles depends on the initial castle assignment to the various armies. Since we are describing an algorithm here, the individual armies have to follow fixed rules of engagement, which are the following.

The exercise starts by distributing the soldiers of the armies into some of the castles and assigning to those castles that have soldiers in them (and to the soldiers themselves) the strength 1 and to all other castles the strength 0. We say that this distribution of armies and strengths describes the situation at the start of Iteration 1. At any given time, a castle will be occupied by the soldiers of the armies that were not weaker than any other soldiers who reached that castle by that time.

The exercise proceeds in discrete iterative steps and the total number of iterations (NI) is determined by the number of distinct affinity values for all armies and roads. These values are put into a strictly decreasing order and the strength of the iteration i ($IS(i)$) is defined as the ith number of this sequence. The following gets done during Iteration i. Those soldiers (and only those soldiers) that occupy a castle of strength $IS(i)$ will try to increase the territory occupied by their army. They will send units from their castle toward all the other castles. When these units arrive at another castle, their strength will be defined as the minimum of $IS(i)$ and the affinity for their army of the road from the originally occupied castle to the new one. If the strength of the new castle is greater than the strength of any of the armies arriving at it, the strength and occupancy of the castle will not change. If no arriving army has greater strength than the strength of the new castle, then the strength of the new castle does not change, but it will get occupied also by those arriving armies whose strength matches that of the castle (but not by any of the others). If some of the arriving armies have greater strength than the strength of the castle, then the castle will be taken over by those (and only those) arriving armies that have the greatest strength, and the strength of the castle is set to the strength of the new occupiers. This describes what happens in one iterative step except for one detail: if an army

gets to occupy a new castle because its strength is $IS(i)$ (this can only happen if the affinity for this army of the road to this castle is at least $IS(i)$), then that army is allowed to send out units from this new castle as well. (This cannot lead to an infinite loop, since there are only finitely many castles and so it can only happen finitely many times that an army gets to occupy a *new* castle because its strength is $IS(i)$.)

The exercise stops at the end of Iteration NI. The output of the exercise provides, for each castle, the strength of the castle and the armies that occupy it at the end of the exercise.

Proof of Theorem 1.1(i). In this existence proof (first published in [28]) we provide an inductive definition that resembles both the description above and the actual algorithm that is described later on. However, this inductive definition is not strictly identical to the algorithm since it was designed to make our proof simple, while the algorithm was designed to be efficient. (In fact, the picturesque description above is nearer to the algorithm than to the inductive definition that follows.)

Let $R = \{1\} \cup \{\psi_m(c,d) > 0 \mid 1 \le m \le M,\ c,d \in V\}$. R is a finite set of real numbers from $(0,1]$, and so its elements can be put into a strictly decreasing order $1 = {}^1r > {}^2r > \cdots > {}^{|R|}r > 0$. We define inductively a sequence of M-semisegmentations ${}^1\sigma,\ {}^2\sigma,\ \ldots,\ {}^{|R|}\sigma$ and a sequence ${}^1U,\ {}^2U,\ \ldots,\ {}^{|R|}U$ of subsets of V as follows.

For any $c \in V$ and $1 \le m \le M$,

$$ {}^1\sigma_m^c = \begin{cases} 1 & \text{if there is a chain of } \psi_m\text{-strength 1 from a seed in } V_m \text{ to } c, \\ 0 & \text{otherwise.} \end{cases} \tag{12.6} $$

(Here, and later, the definition of ${}^i\sigma_0^c$ implicitly follows from the fact that ${}^i\sigma$ is an M-semisegmentation.) For $1 \le i \le |R|$, we define

$$ {}^iU = \left\{ c \mid {}^i\sigma_0^c \ge {}^ir \right\}. \tag{12.7} $$

For $1 < i \le |R|$, $c \in V$, and $1 \le m \le M$, we define

$$ {}^i\sigma_m^c = \begin{cases} {}^{(i-1)}\sigma_m^c & \text{if } c \in {}^{(i-1)}U, \\ {}^ir & \text{if there is a chain } \left\langle c^{(0)}, \ldots, c^{(K)} \right\rangle \text{ of } \psi_m\text{-strength} \\ & {}^ir \text{ such that } c^{(0)} \in {}^{(i-1)}U,\ {}^{(i-1)}\sigma_m^{c^{(0)}} > 0, \\ & c^{(K)} = c \text{ and, for } 1 \le k \le K,\ c^{(k)} \notin {}^{(i-1)}U, \\ 0 & \text{otherwise.} \end{cases} \tag{12.8} $$

It is obvious from these definitions that ${}^{i}\sigma$ is an M-semisegmentation, for $1 \leq i \leq |R|$.

We now demonstrate the definitions on the seeded two-fuzzy graph $\left(\overline{V}, \overline{\Psi}, \overline{\mathcal{V}}\right)$ discussed above. In this case $R = \{1,\ 0.5,\ 0.25\}$. It immediately follows from Eq. (12.6) that ${}^{1}\sigma^{(-1)} = (1,\ 0,\ 1)$, ${}^{1}\sigma^{(0)} = (1,\ 1,\ 0)$, and ${}^{1}\sigma^{(1)} = (0,\ 0,\ 0)$. It turns out that ${}^{2}\sigma = {}^{1}\sigma$. This is because ${}^{1}U = \{(-1),\ (0)\}$, and there are no chains starting at either of these spels which satisfy, for $i = 2$, all the conditions listed in the second line of Eq. (12.8). On the other hand, the chain $\langle (0),\ (1) \rangle$ can be used to generate ${}^{3}\sigma$, which is in fact the 2-segmentation specified by the condition of Theorem 1.1. This is not an accident, we are now going to prove that in general the ${}^{|R|}\sigma$ defined by Eq. (12.6)–(12.8) satisfies the property stated in Theorem 1.1(i).

It clearly follows from the definitions (12.6) and (12.8) that, for $c \in V$ and $1 \leq m \leq M$, ${}^{|R|}\sigma_m^c \in R \cup \{0\}$. Furthermore, it is also not difficult to see, for $1 < i \leq |R|$, that if $c \in {}^{i}U$, then ${}^{i}\sigma_m^c = {}^{|R|}\sigma_m^c$, and that

$$ {}^{i}U = \left\{ c \,|\, {}^{|R|}\sigma_0^c \geq {}^{i}r \right\}. \tag{12.9} $$

These imply the following two properties of the M-semisegmentation ${}^{|R|}\sigma$.

(A) For $c \in V$ and $1 \leq m \leq M$, ${}^{|R|}\sigma_m^c = 1$ if, and only if, there is a chain of ψ_m-strength 1 from a seed in V_m to c.

(B) For $c \in V$, $1 \leq m \leq M$, and $2 \leq i \leq |R|$, ${}^{|R|}\sigma_m^c = {}^{i}r$ if, and only if, there is a chain $\left\langle c^{(0)}, \cdots, c^{(K)} \right\rangle$ of ψ_m-strength ${}^{i}r$ such that $c^{(0)} \in {}^{(i-1)}U$, ${}^{|R|}\sigma_m^{c^{(0)}} > 0$, $c^{(K)} = c$ and, for $1 \leq k \leq K$, $c^{(k)} \notin {}^{(i-1)}U$.

Let $c, d \in V$. We say that (c, d) is *consistent* if, for $1 \leq m \leq M$, ${}^{|R|}\sigma_m^c = {}^{|R|}\sigma_0^c$ implies that one of the following is true:

$$ c = d; \tag{12.10} $$

$$ {}^{|R|}\sigma_0^d > \min\left({}^{|R|}\sigma_0^c,\ \psi_m(c, d) \right); \tag{12.11} $$

$$ {}^{|R|}\sigma_0^d = \min\left({}^{|R|}\sigma_0^c,\ \psi_m(c, d) \right) \quad \text{and} \quad {}^{|R|}\sigma_m^d = {}^{|R|}\sigma_0^d. \tag{12.12} $$

We now show that, for all $c, d \in V$, (c, d) is consistent.

To do this, we assume that there is a (c, d) and an m such that ${}^{|R|}\sigma_m^c = {}^{|R|}\sigma_0^c$ and yet none of Eqs. (12.10)–(12.12) holds and show that this leads to a contradiction. A consequence of our assumption is that $c \neq d$ and at least one of the

following must be the case:

$$ {}^{|R|}\sigma_0^d < \min\left({}^{|R|}\sigma_0^c,\ \psi_m(c,d)\right); \tag{12.13} $$

$$ {}^{|R|}\sigma_0^d = \min\left({}^{|R|}\sigma_0^c,\ \psi_m(c,d)\right) \quad \text{and} \quad {}^{|R|}\sigma_m^d \neq {}^{|R|}\sigma_0^d. \tag{12.14} $$

We may assume that ${}^{|R|}\sigma_0^c > 0$ and that $\psi_m(c,d) > 0$, for otherwise one of Eqs. (12.11) or (12.12) clearly holds. Hence ${}^{|R|}\sigma_m^c = {}^{|R|}\sigma_0^c = {}^{i}r$, for some $1 \leq i \leq |R|$. From Eqs. (12.13) and (12.14) it follows that ${}^{|R|}\sigma_0^d \leq {}^{i}r$. It then follows from Eq. (12.9) that if $i \geq 2$, then neither c nor d is in ${}^{(i-1)}U$.

If $i = 1$, then by A there is a chain of ψ_m-strength 1 from a seed in V_m to c. If $i \geq 2$, then by B there is a chain $\left\langle c^{(0)}, \ldots, c^{(K)} \right\rangle$ of ψ_m-strength ${}^{i}r$ such that $c^{(0)} \in {}^{(i-1)}U$, ${}^{|R|}\sigma_m^{c^{(0)}} > 0$, $c^{(K)} = c$, and, for $1 \leq k \leq K$, $c^{(k)} \notin {}^{(i-1)}U$. In both cases, either d is already in the chain or we can extend the chain without losing its just stated property to d, and so A or B implies that ${}^{|R|}\sigma_m^d = {}^{i}r$. It follows that if $\psi_m(c,d) \geq {}^{i}r$ Eq. (12.12) holds, a contradiction. So assume that $\psi_m(c,d) = {}^{j}r$ for some $j > i$. Since Eq. (12.13) or (12.14) holds, we get that $d \notin {}^{(j-1)}U$. But $c \in {}^{(j-1)}U$, and so, applying B to the chain $\langle c, d\rangle$, we get that ${}^{|R|}\sigma_m^d = {}^{j}r$. This implies that Eq. (12.12) holds. This final contradiction completes our proof that, for all $c, d \in V$, (c,d) is consistent.

Next we show that, for all $c \in V$ and $1 \leq m \leq M$,

$$ {}^{|R|}\sigma_m^c = \mu_{{}^{|R|}\sigma, m, V_m}(c). \tag{12.15} $$

To simplify the notation, we use in this proof s to abbreviate ${}^{|R|}\sigma_m^c$. Recall that $\mu_{{}^{|R|}\sigma, m, V_m}(c)$ denotes the maximal ψ_m-strength of an ${}^{|R|}\sigma m$-chain from a seed in V_m to c. Note that we can assume that $s \in R$, for the alternative is that $s = 0$ in which case there can be no ${}^{|R|}\sigma m$-chain that includes c and so the right-hand side of Eq. (12.15) is also 0 by definition. Our proof will be in two stages: first we show that there is an ${}^{|R|}\sigma m$-chain from a seed in V_m to c of ψ_m-strength s and then we show that there is no ${}^{|R|}\sigma m$-chain from a seed in V_m to c of ψ_m-strength greater than s.

To show the existence of an ${}^{|R|}\sigma m$-chain from a seed in V_m to c of ψ_m-strength s, we use an inductive argument. If $s = {}^{1}r = 1$, then the desired result is assured by A. Now let $i > 1$ and $s = {}^{i}r$. Assume that, for $1 \leq j < i$, whenever a spel d is such that ${}^{|R|}\sigma_m^d = {}^{j}r$, then there is an ${}^{|R|}\sigma m$-chain from a seed in V_m to d of ψ_m-strength ${}^{j}r$.

By B there is a chain $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ of ψ_m-strength s such that $c^{(0)} \in {}^{(i-1)}U$, ${}^{|R|}\sigma_m^{c^{(0)}} > 0$, $c^{(K)} = c$, and, for $1 \leq k \leq K$, $c^{(k)} \notin {}^{(i-1)}U$. We are now going to show that $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ is an ${}^{|R|}\sigma m$-chain by showing that, for $1 \leq k \leq K$, ${}^{|R|}\sigma_m^{c^{(k)}} = s$. Otherwise, consider the smallest $k \geq 1$ that violates this equation. Then we have that ${}^{|R|}\sigma_m^{c^{(k-1)}} \geq s$ and ${}^{|R|}\sigma_m^{c^{(k)}} < s$ (recall that $c^{(k)} \notin {}^{(i-1)}U$ and so ${}^{|R|}\sigma_0^{c^{(k)}} \leq s$). This combined with the fact that $\psi_m\left(c^{(k-1)}, c^{(k)}\right) \geq s$ violates the consistency of $\left(c^{(k-1)}, c^{(k)}\right)$. Since $c^{(0)} \in {}^{(i-1)}U$ and ${}^{|R|}\sigma_m^{c^{(0)}} > 0$, ${}^{|R|}\sigma_m^{c^{(0)}} = {}^{j}r$ for some $1 \leq j < i$ and, by the induction hypothesis, there is an ${}^{|R|}\sigma m$-chain from a seed in V_m to $c^{(0)}$ of ψ_m-strength ${}^{j}r > s$. Appending $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ to this chain we obtain ${}^{|R|}\sigma m$-chain from a seed in V_m to c of ψ_m-strength s. (Just appending may not result in a chain, since a chain is defined as a sequence of distinct spels. However, this is easily remedied by removing, for a repeated spel in the sequence, all the spels between the repetitions and one of the repetitions.)

Now we show that there is no ${}^{|R|}\sigma m$-chain from a seed in V_m to c of ψ_m-strength greater than s. This is clearly so if $s = 1$. Suppose now that $s < 1$ and that $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ is an ${}^{|R|}\sigma m$-chain from a seed in V_m of ψ_m-strength $t > s$. We now show that, for $0 \leq k \leq K$, ${}^{|R|}\sigma_m^{c^{(k)}} \geq t$. From this it follows that $c^{(K)}$ cannot be c and we are done. Since $c^{(0)}$ is a seed in V_m, ${}^{|R|}\sigma_m^{c^{(0)}} = 1 \geq t$. For $k > 0$, induction that makes use of the consistency of $\left(c^{(k-1)}, c^{(k)}\right)$ leads to the desired result.

To show that $\sigma = {}^{|R|}\sigma$ satisfies the property stated in Theorem 1.1(i), we first make two preliminary observations:

(A) For any $c \in V$ and $1 \leq m \leq M$, if $\sigma_n^c > 0$, then $s_n^c = \sigma_n^c = \sigma_0^c$. (The first equality follows from Eqs. (12.14) and (12.15), and the second from the definition of M-semisegmentation.)

(B) For any $c \in V$ and $1 \leq n \leq M$, if $\sigma_n^c = 0$ and $\sigma_0^c > 0$, then $s_n^c < \sigma_0^c$. (Assume the contrary. It cannot be that s_n^c is defined by the first line of Eq. (12.4), for then $c \in V_n$ and by **A** we would have that $\sigma_n^c = 1$. Hence s_n^c is defined by the second line of Eq. (12.4) using some d such that $\min\left(\mu_{\sigma,n,V_n}(d), \psi(d,c)\right) = s_n^c \geq \sigma_0^c > 0$. Hence, by Eq. (12.15) $\sigma_n^d \geq \sigma_0^c > 0$ and so $\sigma_0^d = \sigma_n^d \geq \sigma_0^c$. Interchanging c and d in the definition of consistency, we see that Eq. (12.10) cannot hold since $\sigma_n^d > 0$ and $\sigma_n^c = 0$, (12.11) cannot hold since $\sigma_0^c \leq \sigma_0^d$, and (12.12) cannot hold since $\sigma_n^c = 0$ and $\sigma_0^c > 0$. This contradiction with the consistency of (d, c) proves **B**.)

To complete the proof, let $c \in V$. We first assume that $\sigma_0^c = 0$. Let $1 \leq n \leq M$. By the definition of M-semisegmentation, $\sigma_n^c = 0$. It follows from **A** that $c \notin V_n$ and so s_n^c is defined by the second line of Eq. (12.4). If s_n^c were greater than 0, then there would have to be a $d \in V$ and a chain of positive ψ_n-strength from V_n to d, such that $\psi_n(d, c) > 0$. But then, that chain to d either contains c or could be extended to a chain of positive ψ_n-strength from V_n to c; either case would imply by Eq. (12.15) that $\sigma_n^c > 0$. Hence $s_n^c = 0$, and since this is true for $1 \leq n \leq M$, Eq. (12.5) holds for $1 \leq m \leq M$.

We now assume that $\sigma_0^c > 0$. By the definition of an M-semisegmentation, for $1 \leq m \leq M$, either $\sigma_n^c = \sigma_0^c$ (and there is at least one such n) or $\sigma_n^c = 0$. In the first case we have, by **A**, that $s_n^c = \sigma_n^c = \sigma_0^c$, and in the second case we have, by **B**, that $s_n^c < \sigma_0^c$. From this it again follows that Eq. (12.5) holds for $1 \leq m \leq M$. □

Next we show that such M-semisegmentation is unique. The following proof was first published in [29].

Proof of Theorem 1.1(ii). Suppose that there are two different M-semisegmentations σ and τ of V having the stated property. We choose a spel c, such that $\sigma^c \neq \tau^c$, but for all $d \in V$ such that $\max(\sigma_0^d, \tau_0^d) > \max(\sigma_0^c, \tau_0^c)$, $\sigma^d = \tau^d$. Without loss of generality, we assume that $\sigma_0^c \geq \tau_0^c$, from which it follows that, for some $m \in \{1, \ldots, M\}$, $\sigma_m^c > \tau_m^c (\geq 0)$ and so, by Eq. (12.5), $\sigma_m^c = s_m^c$ and $c \notin V_m$. This implies that there exists a σm-chain $\langle d^{(0)}, \ldots, d^{(L)} \rangle$ in V of ψ_m-strength not less than $\sigma_m^c (> 0)$ such that $d^{(0)} \in V_m$ and $\psi_m(d^{(L)}, c) \geq \sigma_m^c$. Next we show that $\langle d^{(0)}, \ldots, d^{(L)} \rangle$ is a τm-chain.

We need to show that for $0 \leq l \leq L$, $\tau_m^{d^{(l)}} > 0$. This is true for 0, since $d^{(0)} \in V_m$. Now assume that it is true for $l - 1$ $(1 \leq l \leq L)$. Since $\langle d^{(0)}, \ldots, d^{(l-1)} \rangle$ is a τm-chain in V of ψ_m-strength at least $\sigma_m^c (> 0)$ from an element of V_m, we have that $\mu_{\tau, m, V_m}(d^{(l-1)}) \geq \sigma_m^c$. Since we also know that $\psi_m(d^{(l-1)}, d^{(l)}) \geq \sigma_m^c$, we get that $t_m^{d^{(l)}} \geq \sigma_m^c$ (where t is defined for τ as s is defined for σ in Eq. (12.4)). The only way $\tau_m^{d^{(l)}}$ could be 0, if there were an $n \in \{1, \ldots, M\}$ such that $\max(\sigma_0^{d^{(l)}}, \tau_0^{d^{(l)}}) \geq \tau_0^{d^{(l)}} = \tau_n^{d^{(l)}} = t_n^{d^{(l)}} > t_m^{d^{(l)}} \geq \sigma_m^c = \sigma_0^c = \max(\sigma_0^c, \tau_0^c)$. By the choice of c, this would imply that $\sigma^{d^{(l)}} = \tau^{d^{(l)}}$, which cannot be since $\sigma_m^{d^{(l)}} \neq 0$.

From the facts that $\langle d^{(0)}, \ldots, d^{(L)} \rangle$ is a τm-chain of ψ_m-strength not less than σ_m^c and that $\psi_m(d^{(L)}, c) \geq \sigma_m^c$, it follows that $\tau_0^c \geq t_m^c \geq \sigma_m^c = \sigma_0^c \geq \tau_0^c$, implying that all the inequalities are in fact equalities. But then $\sigma_m^c = t_m^c = \tau_m^c$, contradicting $\sigma_m^c > \tau_m^c$ and thereby validating uniqueness. □

Finally we show that provided that $(V, \Psi, \mathcal{V})$ is connectable, any M-semisegmentation having the stated property is in fact an M-segmentation. The following proof was also first published in [29].

Proof of Theorem 1.1(iii). We observe that it is a consequence of Eq. (12.5) that, for any spel c, $\sigma_0^c = \max_{1 \le m \le M} s_m^c$. Since we assume that the seeded M-fuzzy graph $(V, \Psi, \mathcal{V})$ is connectable, there exists a chain $\langle c^{(0)}, \ldots, c^{(K)} \rangle$ of positive $(\min_{1 \le m \le M} \psi_m)$-strength from a seed spel to an arbitrary spel c. We now show inductively that, for $0 \le k \le K$, $\sigma_0^{c^{(k)}} > 0$. This is clearly so for $k = 0$. Suppose now that it is so for $k - 1$. Choose an m $(1 \le m \le M)$ such that $\sigma_0^{c^{(k-1)}} = \sigma_m^{c^{(k-1)}} = s_m^{c^{(k-1)}}$. Then there is a σm-chain of positive ψ_m-strength from a spel in V_m to $c^{(k-1)}$. Since $\psi_m(c^{(k-1)}, c^{(k)}) > 0$, $\sigma_0^{c^{(k)}} \ge s_m^{c^{(k)}} > 0$. □

12.3.2 Algorithm

In this subsection, we present an algorithm that produces the M-semisegmentations whose existence and uniqueness are guaranteed by Theorem 1.1. In designing the algorithm we aimed at making it efficient: as is illustrated in the next subsection, our implementation of it allowed us to find 3-segmentations of images with over 10,000 spels in approximately a tenth of a second.

As the algorithm proceeds, it maintains (and repeatedly changes) an M-semisegmentation σ. The claim is that at the time when the algorithm terminates, σ satisfies the property of Theorem 1.1(i).

The algorithm makes use of a priority queue H of spels c, with associated keys σ_0^c [26]. Such a priority queue has the property that the key of the spel at its head is maximal (its value is denoted by Maximum-Key(H), which is defined to be 0 if H is empty). As the algorithm proceeds, each spel is inserted into H exactly once (using the operation $H \leftarrow H \cup \{c\}$) and is eventually removed from H using the operation Remove-Max(H), which removes the spel c from the head of the priority queue. At the time when a spel c is removed from H, the vector σ^c has its final value. Spels are removed from H in a nonincreasing order of the final value of σ_0^c. We use the variable l to store the current value of Maximum-Key(H). Algorithm 1 shows a detailed specification using the conventions adopted in [26].

The process is initialized (Steps 1–10) by first setting σ_m^c to 0, for each spel c and $0 \le m \le M$. Then, for every seed spel $c \in V_m$, c is put into U_m and H and

Algorithm 1 Multiobject fuzzy segmentation.

```
1.  for c ∈ V
2.     do for m ← 0 to M
3.        do σ_m^c ← 0
4.  H ← ∅
5.  for m ← 1 to M
6.     do U_m ← V_m
7.        for c ∈ U_m
8.           do if σ_0^c = 0 then H ← H ∪ {c}
9.              σ_0^c ← σ_m^c ← 1
10. l ← 1
11. while l > 0
12.    for m ← 1 to M
13.       do while U_m ≠ ∅
14.          do remove a spel d from U_m
15.             C ← {c ∈ V | σ_m^c < min(l, ψ_m(d, c)) and
                          σ_0^c ≤ min(l, ψ_m(d, c))}
16.             while C ≠ ∅
17.                do remove a spel c from C
18.                   t ← min(l, ψ_m(d, c))
19.                   if l = t and σ_m^c < l then U_m ← U_m ∪ {c}
20.                   if σ_0^c < t then
21.                      if σ_0^c = 0 then H ← H ∪ {c}
22.                      for n ← 1 to M
23.                         do σ_n^c ← 0
24.                   σ_0^c ← σ_m^c ← t
25.    while Maximum-Key(H) = l
26.       Remove-Max(H)
27.    l ← Maximum-Key(H)
28.    for m ← 1 to M
29.       U_m ← {c ∈ H | σ_m^c = l}
```

both σ_0^c and σ_m^c are set to 1. Following this, l is also set to 1. At the end of the initialization, the following conditions are satisfied.

(i) σ is an M-semisegmentation of V.

(ii) A spel c is in H if, and only if, $0 < \sigma_0^c \leq l$.

(iii) $l =$ Maximum-Key(H).

(iv) For $1 \leq m \leq M$, $U_m = \{c \in H \mid \sigma_m^c = l\}$.

The initialization is followed by the main loop of the algorithm. At the beginning of each execution of this loop, conditions (i) to (iv) above are satisfied. The main loop is repeatedly performed for decreasing values of l until l becomes 0, at which time the algorithm terminates (Step 11). There are two parts to the main loop, each of which has a very different function.

The first part of the main loop (Steps 12–24) is the essential part of the algorithm. It is in here where we update our best guess so far of the final values of the σ_m^c. A current value is replaced by a larger one if it is found that there is a σm-chain from a seed spel in V_m to c of ψ_m-strength greater than the old value (the previously maximal ψ_m-strength of the known σm-chains of this kind) and it is replaced by 0 if it is found that (for an $n \neq m$) there is a σn-chain from a seed spel in V_n to c of ψ_n-strength greater than the old value of σ_m^c.

The purpose of the second part of the main loop (Steps 25–29) is to restore the satisfaction of conditions (iii) and (iv) above for a new (smaller) value of l.

To help with the understanding of why this algorithm performs as desired, we comment that just prior to entering its main loop (Steps 11–29), there are four kinds of spels. There are those spels d that have previously been put into and have subsequently been removed from H; for these spels not only does the vector σ^d has its final value, but also we have already put into H (and possibly even have already removed from H) every spel c such that $\psi_m(d, c) > 0$, for at least one m. (For spels of this first kind, $\sigma_0^d > l$.) Secondly, there are the spels d that are in at least one of the U_m; for these spels the vector σ^d has its final value, but we may not have yet put into H every spel c such that $\psi_m(d, c) > 0$ for at least one m. (For spels of this second kind, $\sigma_0^d = \sigma_m^d = l$.) This will get done in the next execution of Steps 13–21, while Steps 22–24 will insure that the σ^c get updated appropriately. Consequently, the spels c which are in H but not in any of the U_m are those for which there is, for some $1 \leq m \leq M$, a σm-chain (for the

current σ) from a seed spel in V_m to c; for the rest of the spels (those which have not as yet been put into H) there is no m for which there is a σm-chain (for the current σ) from a seed spel in V_m to c. (For spels c of these third and fourth kinds, $0 < \sigma_0^c < l$ and $\sigma_0^c = 0$, respectively.)

One tricky aspect of the algorithm is that a spel of the third kind may become a spel of the second kind and a spel of the fourth kind may become a spel of the third (or even of the second) kind during the execution of the main loop. That the description of the four kinds of spels remains as given in the previous paragraph is insured by Steps 19 and 21. (Step 21 also insures that condition (ii) stated above remains satisfied. To see this, observe that Step 15 guarantees that if c is put into C, then $0 < \min(l, \psi_m(d, c))$ and consequently the t defined in Step 18 and used in Step 24 is also positive. That condition (i) stated above remains satisfied is obvious from Steps 20–24.)

We complete this subsection with a brief discussion of our implementation of Algorithm 1. As suggested in [26], we use a heap to implement the priority queue H. This provides us with efficient implementations of the operations of insertion into ($H \leftarrow H \cup c$) and removal from (Remove-Max(H)) the priority queue, as well as of Step 29. In applications it is typically the case that, for every spel d, there is a fixed number of spels c such that $\Sigma_{m=1}^{M} \psi_m(d, c) > 0$ and a list of all such c is inexpensive to produce. In such a case the cost of executing Step 15 becomes proportional to a constant (four, six, or twelve in the examples shown below and in Section 12.4) independent of the size of V. Using L to denote this constant, the computational complexity of the Algorithm 1 is the following: since each spel can belong to multiple objects there can be at most $M|V|$ executions of the loop 13–24, while the loop 16–24 can be executed at most L times. Steps 19 and 24 have $O(\log |V|)$ operations while Steps 22–23 have $O(M)$ operations, so the loop 16–24 has $O(M \log |V|)$ operations. Since this loop can be executed at most $ML|V|$ times, the time complexity of Algorithm 1 is $O(M^2 L|V| \log |V|)$.

12.3.3 Experiments

Now we demonstrate the usage of Algorithm 1 on mathematically-defined as well as on real images. Similarly to the example shown in section 12.2, the appropriate fuzzy spel affinities were automatically defined by a computer program, based on some minimal information supplied by a user. However, this is not the only option: for example, if sufficient prior knowledge about the class of

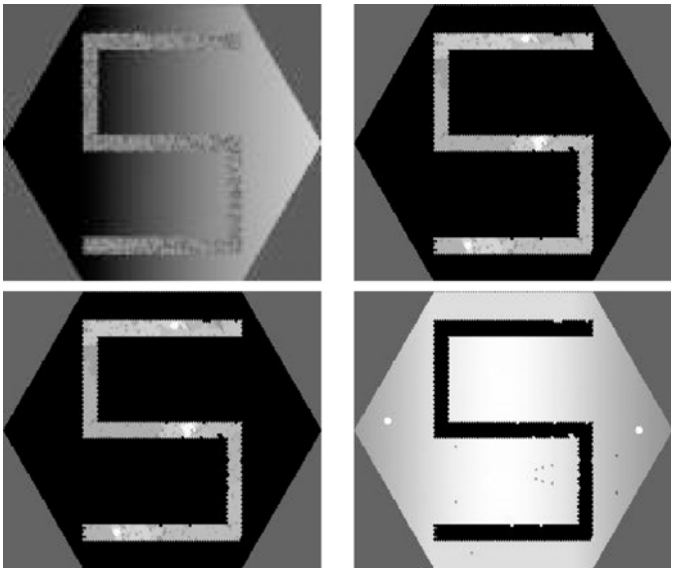


Figure 12.3: A mathematically defined image (top left) including both background variation and noise, and the corresponding 3-segmentation (top right and bottom row).

segmentation problems to which the application at hand belongs is available, then the whole segmentation process can be automated by designing a program that automatically selects the seeds for the objects to be segmented, as it was done in [30] to segment macromolecules in electron microscopy volumes.

On the top-left of Figs. 12.3–12.7 and in the left column of Fig. 12.8 are images defined on a V consisting of regular hexagons that are inside a large hexagon (with 60 spels on each side, a total of 10,621 spels). In all these examples, $M = 3$. For these experiments we defined $V_m\,(1 \leq m \leq 3)$ to be the set of spels indicated by the user plus their six neighbors. The fuzzy affinity functions $\psi_m\,(1 \leq m \leq 3)$ were computed according to Eqs. (12.2) and (12.3), with adjacency π between hexagons meaning that they share an edge.

The other three images of Figs. 12.3–12.7 represent the resulting σ_m (obtained by Algorithm 1) with the brightness of each spel encoding its grade of membership in an object. (For Fig. 12.3 we selected the seed spels so that $V_1 = V_2$, for Fig. 12.4 we selected the seed spels so that $V_2 = V_3$, and for Figs. 12.5–12.7 the three sets of seed spels are pairwise disjoint, which happens to result, because of the large number of gray levels used in the images

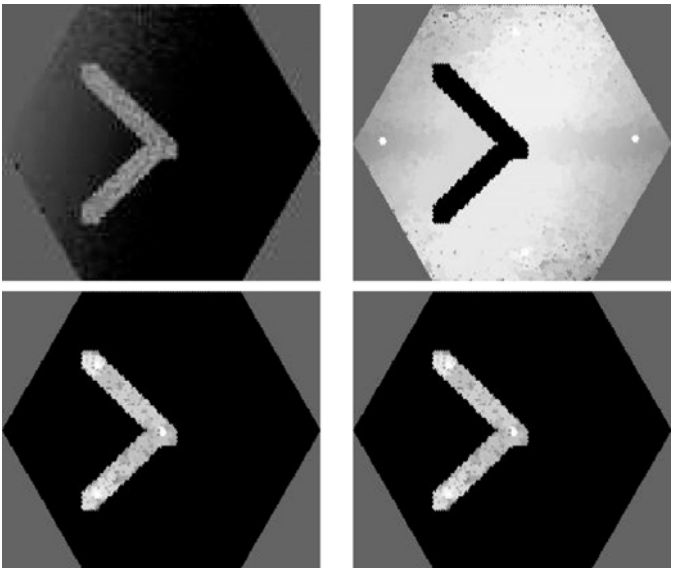


Figure 12.4: A mathematically defined image (top left) including both background variation and noise, and the corresponding 3-segmentation (top right and bottom row).

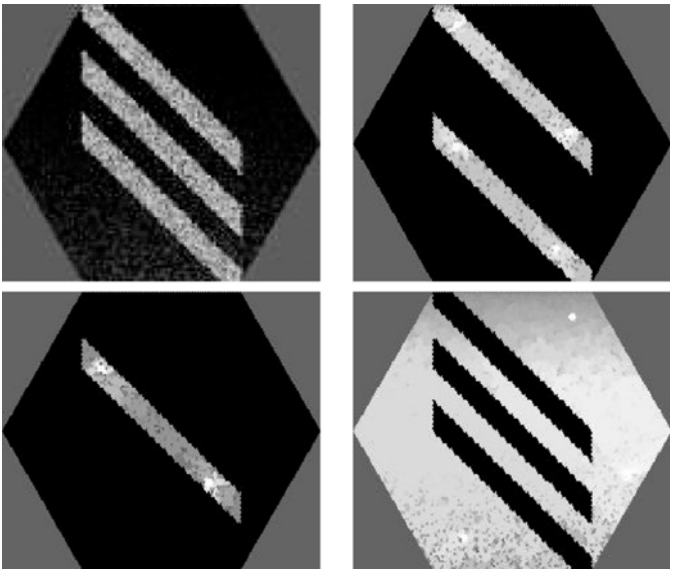


Figure 12.5: A mathematically defined image (top left) including both background variation and noise, and the corresponding 3-segmentation (top right and bottom row).

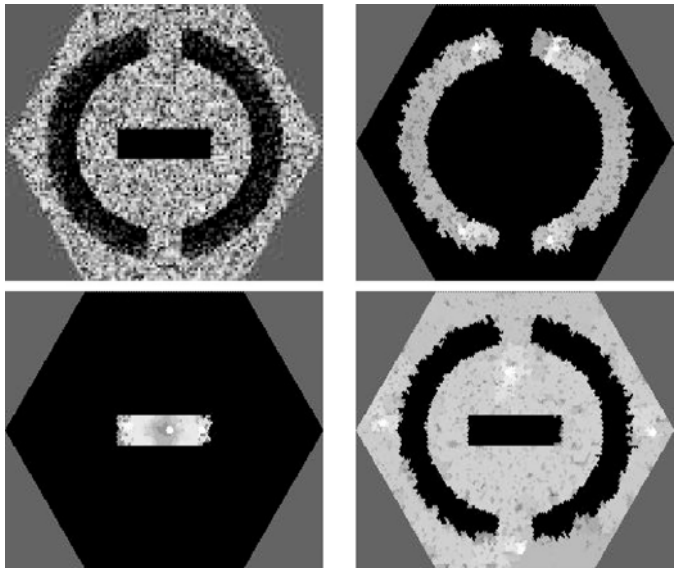


Figure 12.6: A mathematically defined image (top left) including both background variation and noise, and the corresponding 3-segmentation (top right and bottom row).

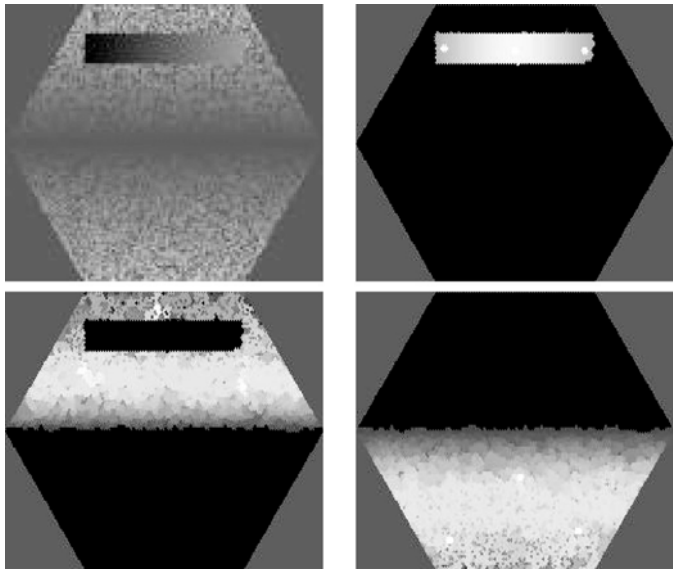


Figure 12.7: A mathematically defined image (top left) including both background variation and noise, and the corresponding 3-segmentation (top right and bottom row).

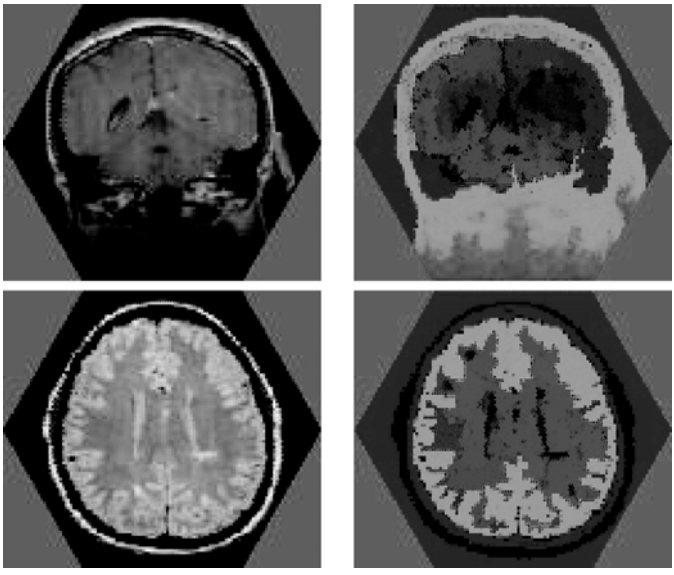


Figure 12.8: Two images obtained using magnetic resonance imaging (MRI) of heads of patients (left) and the corresponding 3-segmentations (right). (Color slide.)

to be segmented, in the three objects being pairwise disjoint as well.) The right column of Fig. 12.8 shows the resulting maps of the σ_m by assigning the color $(r, g, b) = 255 \times (\sigma_1^c, \sigma_2^c, \sigma_3^c)$ to the spel c. Note that not only the hue, but also the brightness of the color is important: the less brightly red areas for the last two images correspond to the ventricular cavities in the brain, correctly reflecting a low grade of membership of these spels in the object that is defined by seed spels in brain tissue. The seed sets V_m consist of the brightest spels. The times taken to calculate these 3-segmentations using our algorithm on a 1.7 GHz Intel® Xeon™ personal computer were between 90 and 100 ms for each of the seven images (average = 95.71 ms). Since these images contain 10,621 spels, the execution time is less than 10 μs per spel. The same was true for all the other 2-D image segmentations that we tried, some of which are reported in what follows.

To show the generality of our algorithm and to permit comparisons with other algorithms, we also applied it to a selection of real images that appeared in the recent image segmentation literature. Since in all these images V consist of squares inside a rectangular region, the π of Eq. (12.2) is selected to

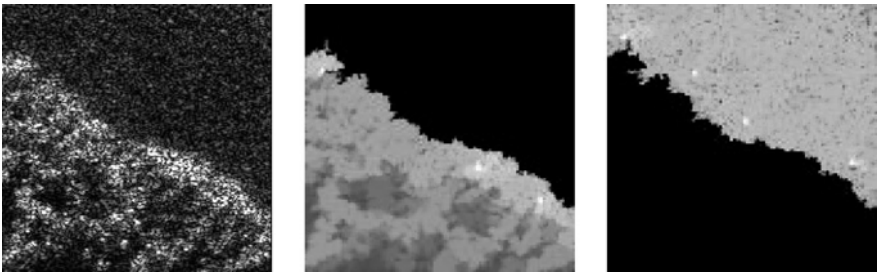


Figure 12.9: An SAR image of trees and grass (left) and its 2-segmentation (center and right).

be the edge-adjacency (4-adjacency) on the square grid. We chose to include in the sets of seed spels not only the spels at which the user points but also their eight edge-or-vertex adjacent spels. Except for this adaptation, the previous specification is verbatim what we use for the experiments which we now describe.

In [31] the authors demonstrate their proposed technique by segmenting an SAR image of trees and grass (their Fig. 1, our Fig. 12.9 left). They point out that "the accurate segmentation of such imagery is quite challenging and in particular cannot be accomplished using standard edge detection algorithms." They validate this claim by demonstrating how the algorithm of [32] fails on this image. As illustrated on the middle and right images of Fig. 12.9, our technique produces a satisfactory segmentation. On this image, the computer time needed by our algorithm was 0.3 s (on the same 1.7 GHz Intel® Xeon™ personal computer that we use for all experiments presented in this section), while according to a personal communication from the first author of [31], its method "took about 50 seconds to reach the 2-region segmentation for this 201-by-201 image on Sparc 20, with the code written in C."

In Figs. 12.10–12.12 we report on the results of applying our approach to two physically obtained images from [6]: an aerial image of San Francisco (top-left image of Fig. 12.10) and an indoor image of a room (top-left image of Fig. 12.11). The middle and bottom images of the left column of Fig. 12.10 show a 2-segmentation of the San Francisco image into land and sea, while the right column shows how extra object definitions can be included in order to produce a more detailed labeling of a scene, with the 3-segmentation of the San Francisco image separating the Golden Gate Park from the rest of the land object. Figure 12.11 shows the original image (top-left) and a 5-segmentation of the

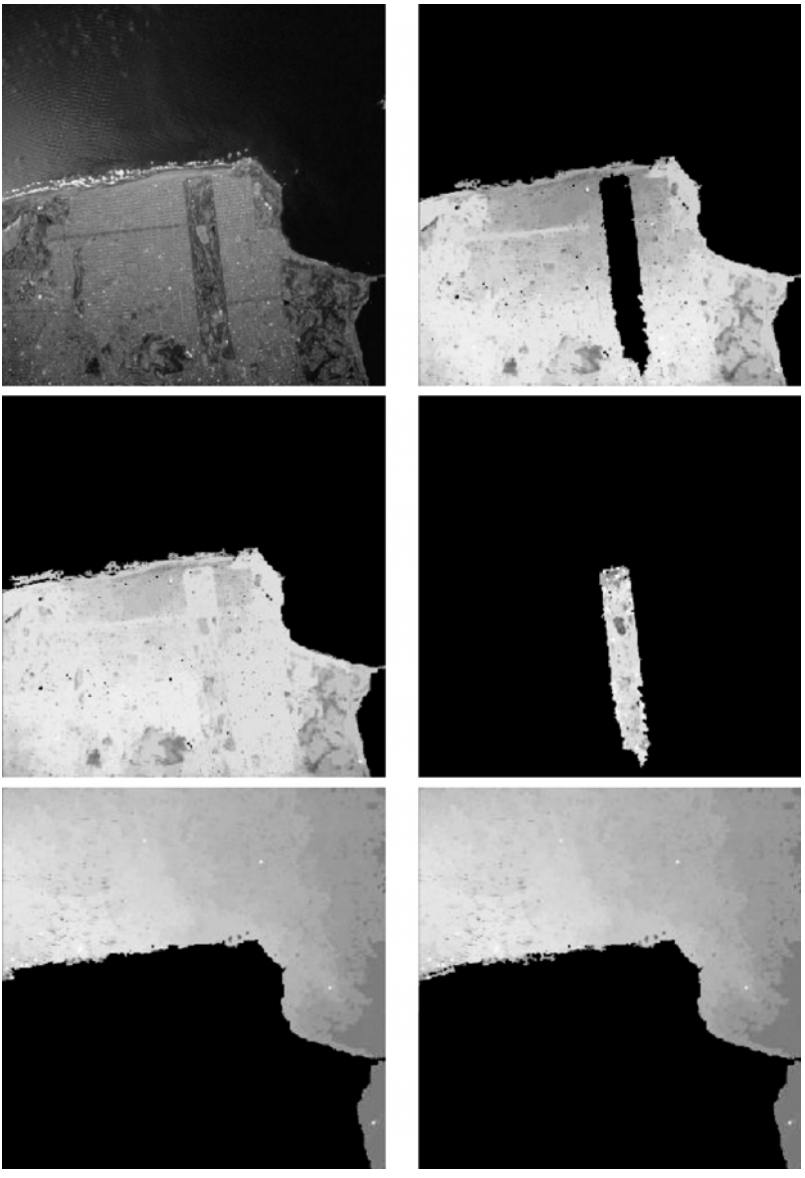


Figure 12.10: Aerial image of San Francisco (top left), a 2-segmentation into land and sea (middle and bottom images on the left column) and a 3-segmentation into built-up land, the Golden Gate Park, and sea (right column).

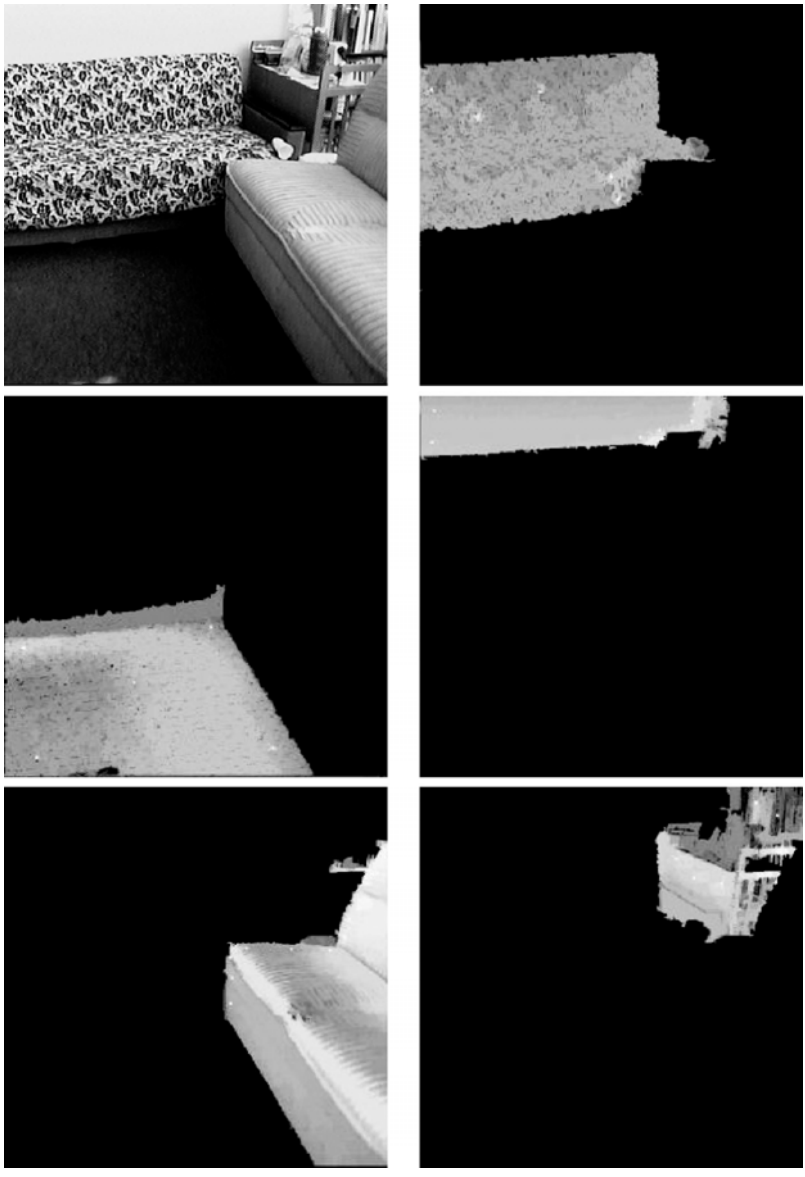


Figure 12.11: An indoor image of a living room (top left) and its 5-segmentation.

living room image. The 6-segmentation of the room shown in Fig. 12.12 includes a new object corresponding to the base and arm of one of the sofas.

It is stated in [6] that the times needed for the segmentations reported in that paper "are in the range of less than five seconds" (on a Sun UltraSparc™). Our CPU time to obtain the segmentations shown in Figs. 12.10–12.12 is

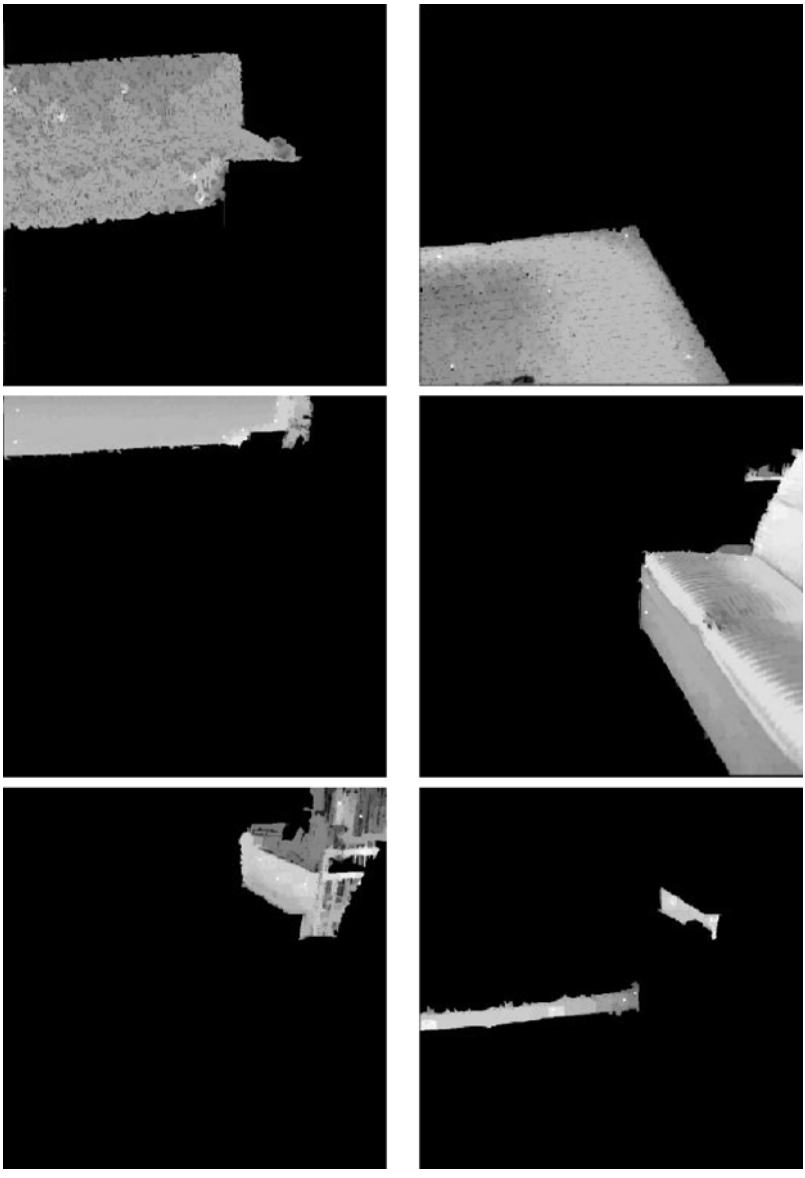


Figure 12.12: A 6-segmentation of the indoor image of a living room shown in Fig. 12.11.

around 2 s. However, there is a basic difference in the resolutions of the segmentations. Since the segmentation method used in [6] is texture based, the original 512×512 images are subdivided into 64×64 "sites" using a square window of size 8×8 per site. In the final segmentations of [6] all pixels in a particular

window are assigned to the same object. As opposed to this, in our segmentations any pixel can be assigned to any object. Another way of putting this is that we could also make our spels to be the 8×8 windows of [6] and thereby reduce the size of the V to be a 64th of what it is currently. This should result in a two order of magnitude speedup in the performance of our segmentation algorithm (at the cost of a loss of resolution in the segmentations to the level used in [6]).

12.3.4 Accuracy and Robustness

Because all affinities (and consequently the segmentations) shown in the last section are based on seeds selected manually by a user, the practical usefulness and performance of the multiseeded fuzzy segmentation algorithm need to be experimentally evaluated, both for accuracy and for robustness.

The experiments used the top-left images of Figs. 12.3–12.7. We chose these images because they were based on mathematically defined objects to which we assigned gray values that were then corrupted by random noise and shading, and so the "correct" segmentations were known to us.

We then asked five users who were not familiar with the images to perform five series of segmentations, where each series consisted of the segmentation of each one of the five images presented in a random order. Since each of the five users performed five series of segmenting the five images, we had at our disposal 125 segmentations that were analyzed in a number of different ways.

First, we analyzed the segmentations concerning their accuracy. We used two reasonable ways of measuring the accuracy of the segmentations: in one we simply consider if the spel is assigned to the correct object, in the other we take into consideration the grade of membership as well. The *point accuracy* of a segmentation is defined as the number of spels correctly identified divided by the total number of spels multiplied by 100. The *membership accuracy* of a segmentation is defined as the sum of the grades of membership of all the spels which are correctly identified divided by the total sum of the grades of membership of all spels in the segmentation multiplied by 100.

The average and the standard deviation of the point accuracy for all 125 segmentations were 97.15 and 4.72, respectively, while the values for their membership accuracy were 97.70 and 3.82. These means and standard deviations are very similar. This is reassuring, since the definitions of both of the accuracies

are somewhat *ad hoc* and so the fact that they yield similar results indicates that the reported figures of merit are not over-sensitive to the precise nature of the definition of accuracy. The slightly larger mean for the membership accuracy is due to the fact that misclassified spels tend to have smaller than average grade of membership values.

The average *error* (defined as "100 less point accuracy") over all segmentations is less than 3%, comparing quite favorably with the state of the art: in [6] the authors report that a "mean segmentation error rate as low as 6.0 percent was obtained."

The robustness of our procedure was defined based on the similarity of two segmentations. The *point similarity* of two segmentations is defined as the number of spels which are assigned to the same object in the two segmentations divided by the total number of spels multiplied by 100. The *membership similarity* of two segmentations is defined as the sum of the grades of memberships (in both segmentations) of all the spels which are assigned to the same object in the two segmentations divided by the total sum of the grades of membership (in both segmentations) of all the spels multiplied by 100. (For both these measures of similarity, identical segmentations will be given the value 100 and two segmentations in which every spel is assigned to a different object will be given the value 0.)

Since each user segmented each image five times, there are 10 possible ways of pairing these segmentations, so we had 50 pairs of segmentations per user and a total of 250 pairs of segmentations. Because the results for point and membership similarity were so similar for every user and image (for detailed information, see [29]) we decided to use only one of them, the point similarity, as our *intrauser consistency* measure. The results are quite satisfactory, with an average intra-user consistency of 96.88 and a 5.56 standard deviation.

In order to report on the consistency between users (*interuser consistency*) we selected, for each user and each image, the *most typical segmentation* by that user of that image. This is defined as that segmentation for which the sum of membership similarities between it and the other four segmentations by that user of that image is maximal. Thus, we obtained five segmentations for each image that were paired between them into 10 pairs, resulting into a total of 50 pairs of segmentations. The average and standard deviation of the interuser consistency (98.71 and 1.55, respectively) were even better than the intrauser

consistency, mainly because the selection of the most typical segmentation for each user eliminated the influence of relatively bad segmentations.

Finally, we did some calculations of the sensitivity of our approach to M (the predetermined number of objects in the image). The distinction between the objects represented in the top right and bottom left images of Fig. 12.5 and between the objects represented in the bottom images of Fig. 12.7 is artificial; the nature of the regions assigned to these objects is the same. The question arises: if we merge these two objects into one do we get a similar 2-segmentation to what would be obtained by merging the seed points associated with the two objects into a single set of seed points and then applying our algorithm? (This is clearly a desirable robustness property of our approach.) The average and standard deviation of the point similarity under object merging for a total of 50 readings by our five users on the top-left images of Figs. 12.5 and 12.7 were 99.33 and 1.52, respectively.

12.4 3-D Segmentation

As showed before, the multiseeded segmentation algorithm is general enough to be applied to images defined on various grids. One has several options for representing a 3D image; in this section, when performing segmentation on 3D images, we choose to represent them on the face-centered cubic (fcc) grid, for reasons that are presented later.

Using $\mathbb{Z}$ for the set of all integers and δ for a positive real number, we can define the *simple cubic* (sc) grid (S_δ), the *face-centered cubic* (fcc) grid (F_δ) and the *body-centered cubic* (bcc) grid (B_δ) as

$$S_\delta = \{(\delta c_1, \delta c_2, \delta c_3) \mid c_1, c_2, c_3 \in \mathbb{Z}\}\,, \tag{12.16}$$

$$F_\delta = \{(\delta c_1, \delta c_2, \delta c_3) \mid c_1, c_2, c_3 \in \mathbb{Z} \text{ and } c_1 + c_2 + c_3 \equiv 0 \,(\mathrm{mod}\, 2)\}\,, \tag{12.17}$$

$$B_\delta = \{(\delta c_1, \delta c_2, \delta c_3) \mid c_1, c_2, c_3 \in \mathbb{Z} \text{ and } c_1 \equiv c_2 \equiv c_3 \,(\mathrm{mod}\, 2)\}\,, \tag{12.18}$$

where δ denotes the grid spacing. From the definitions above, the fcc and bcc grids can be seen either as one sc grid without some of its grid points or as a union of shifted sc grids, four in the case of the fcc and two in the case of the bcc.

We now generalize the notion of a voxel to an arbitrary grid. Let G be any set of points in $\mathbb{R}^3$, then the *Voronoi neighborhood* of an element g of G is

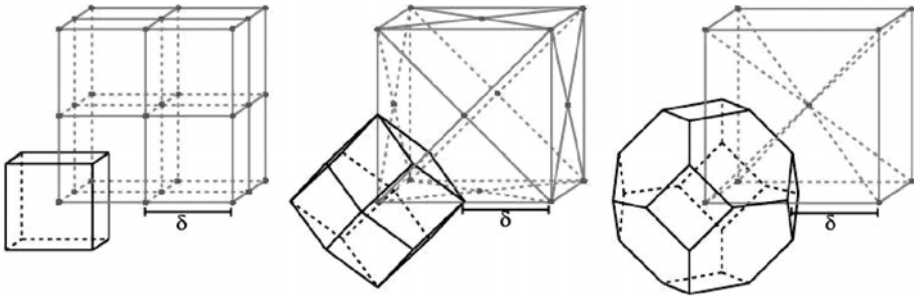


Figure 12.13: Three grids with the Voronoi neighborhood of one of their grid points. From left to right: the simple cubic (sc) grid, the face-centered cubic (fcc) grid, and the body-centered cubic (bcc) grid.

defined as

$$N_G(g) = \{v \in \mathbb{R}^3 \mid \text{for all } h \in G, \ \ \|v - g\| \leq \|v - h\|\}. \tag{12.19}$$

In Fig. 12.13, we can see the sc, the fcc and the bcc grids and the Voronoi neighborhoods of the front-lower-left grid points.

Why should one choose grids other than the ubiquitous simple cubic grid? The fcc and bcc grids are superior to the sc grid because they sample the 3-D space more efficiently, with the bcc being the most efficient of the three. This means that both the bcc and the fcc grid can represent a 3-D image with the same accuracy as that of the sc grid but using fewer grid points [33].

We decided to use the fcc grid for 3-D images instead of the bcc grid for reasons that will become clear in a moment; now we discuss one additional advantage of using the fcc grid over the sc grid. If we have an object that is a union of Voronoi neighborhoods of the fcc grid, then for any two faces on the boundary between this object and the background that share an edge, the normals of these faces make an angle of 60° with each other. This results in a less blocky image than if we used a surface based on the cubic grid with voxels of the same size. This can be seen in Fig. 12.14, where we display approximations to a sphere based on different grids. Note that the display based on the fcc grid (center) has a better representation than the one based on the sc grid with the same voxel volume (left) and is comparable with the representation based on cubic grid with voxel volume equal to one eighth of the fcc voxel volume (right).

The main advantage of the bcc grid over the fcc grid is that it needs fewer grid points to represent an image with the same accuracy. However, in the bcc

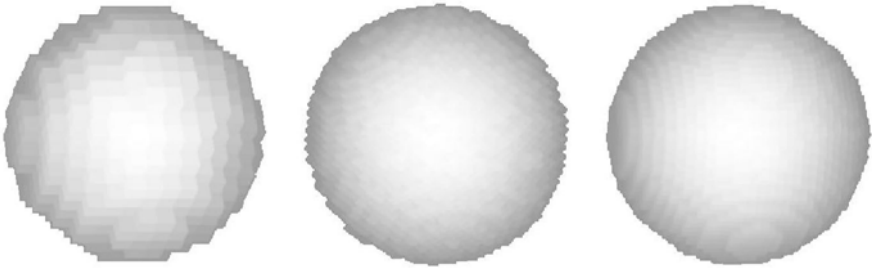


Figure 12.14: Computer graphic display of a sphere using different grids (reproduced from [19]). (a) Display based on a sc grid with voxels of the same volume as the display based on a fcc used for (b). The image (c) corresponds to a display based on a sc grid with voxels of volume equal to one eighth of the voxel volume in the other two images.

grid, grid points whose Voronoi neighborhoods share a face can be at one of two distances from each other, depending on the kind of face they share (see Fig. 12.13), a characteristic that may not be desirable. The Voronoi neighborhoods of an fcc grid F_δ are rhombic dodecahedra (polyhedra with 12 identical rhombic faces), as can be seen in Fig. 12.13. We can define the adjacency relation β for the grid F_δ by: for any pair (c, d) of grid points in F_δ,

$$(c, d) \in \beta \Leftrightarrow \|c - d\| = \sqrt{2}\delta. \tag{12.20}$$

Each grid point $c \in F_\delta$ has 12 β-adjacent grid points in F_δ. In fact, two grid points in F_δ are adjacent if, and only if, the associated Voronoi neighborhood share one face. In practice these definitions give rise to a digital space (V, π) by using a V that is a finite subset of F_δ and a π that is the β of Eq. (12.20) restricted to V. (Note that since Eq. (12.2) ignores distance, a similar approach applied to the bcc grid would have the undesirable consequence of having a fuzzy spel affinity that does not incorporate the difference in distances between adjacent spels.)

Experiments with segmentations using this approach on 3-D images were reported in [34]. Here we present two more recent experiments from [35] of multiple object fuzzy segmentation of 3-D images on the fcc grid.

The first experiment was performed on a computerized tomography (CT) reconstruction that assigned values to a total of $(298 \times 298 \times 164)/2 = 7{,}281{,}928$ (see Eq. 12.17)) fcc grid points. We selected seeds for four objects, the intestine (red object), other soft tissues (green object), the bones (blue object) and the

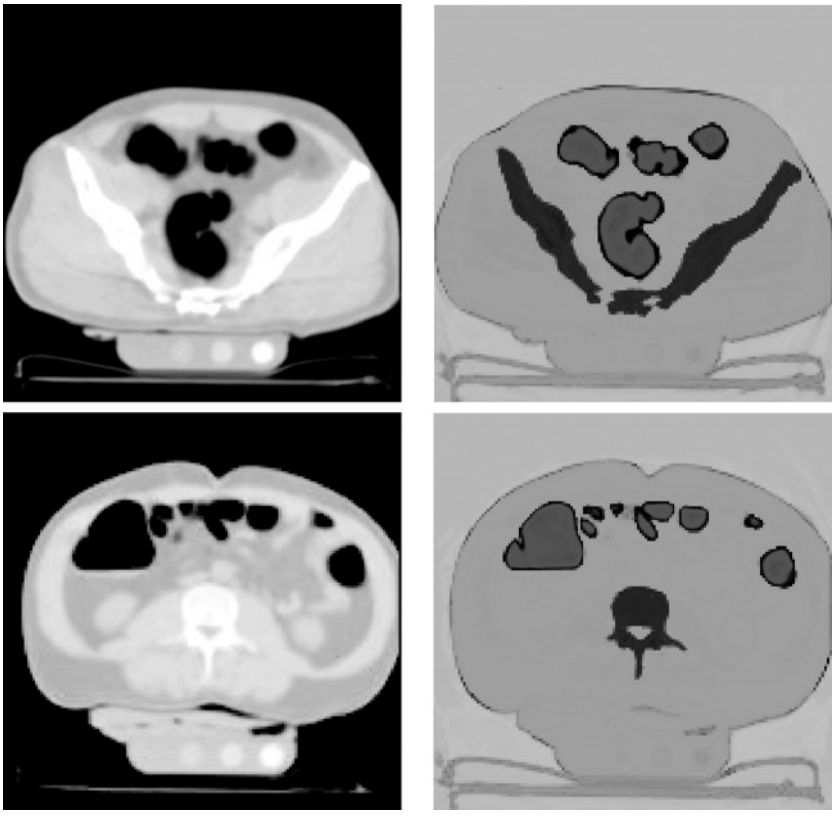


Figure 12.15: Two axial slices from a CT volume placed on the fcc grid and the corresponding 4-segmentations. (All four images were interpolated to the sc grid for display purposes.)

lungs/background (cyan object). Then, using a 1.7 GHz Intel® Xeon™ personal computer, our program performed the 4-segmentation on this volume that is shown in Fig. 12.15.

The execution time of our program was 249 s, or approximately 34 μs were needed per spel to perform the segmentation. Based on the execution timings for the previous experiments, one should expect a smaller execution time for this volume. There are three main reasons why the average number of spels segmented per second is not higher. First, since we have used the β-adjacency, the number of neighboring spels was doubled or tripled when compared to the previous examples, where we used the edge-adjacency for the images on the hexagonal (six neighbors) and square (four neighbors) grids, respectively.

Second, the memory saving approach used in implementing the 3-D version of our algorithm slowed down the execution. (Since the goal was to segment volumes that could have as many as $512 \times 512 \times 512$ spels, we chose to implement a "growing" heap, where a new level of the heap was added or an old one was removed as the program was executed depending on the number of spels currently in the heap, so that the memory usage was kept as low as possible. Note that, besides the heap, both the M-segmentation map and the original volume are accessed simultaneously by Algorithm 1). Finally, our program was developed with the goal of being able to segment images placed on the sc, fcc, and bcc grids, and this generality also contributed to the longer execution time of the algorithm; as opposed to the approach taken in the 2-D case, where we use two programs to produce the segmentations shown in subsection 12.3.3 (one for the images on the hexagonal grid and another for the images on the square grid).

In order to have a better idea of the quality of the segmentation produced by our algorithm on this volume, we created a 3-D model of the segmented intestines (the red object of Fig. 12.15) using the software OpenDX [36], which can be seen in Fig. 12.16. Since OpenDX can work with arbitrary grids, we used the fcc grid: the surface shown on Fig. 12.16 consists of faces of rhombic dodecahedra (fcc Voronoi neighborhoods).

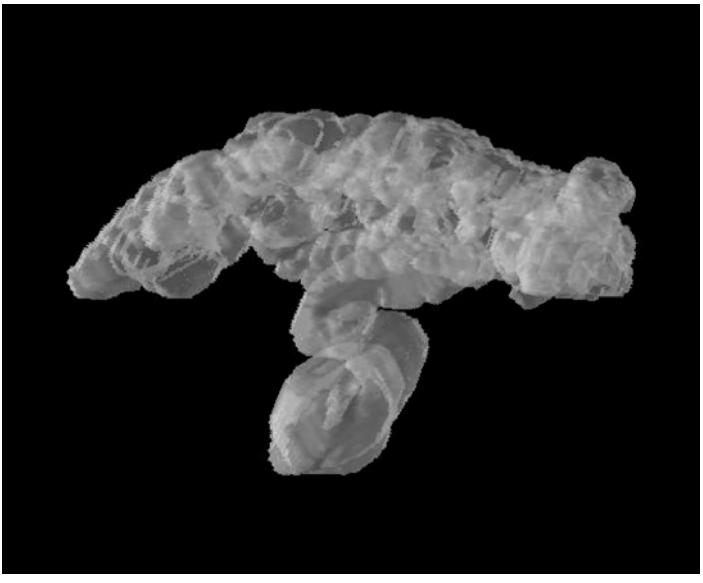


Figure 12.16: A 3-D view of the segmented intestines shown in Fig. 12.15.

For a second experiment, we interpolated a clinically obtained CT dataset from the sc grid to the fcc grid. In this experiment, the aim is to segment the trachea and bronchial tubes so that they can be subsequently visualized in a virtual bronchoscopy (VB) animation [35]. (We refer to this dataset as the VB dataset.) This dataset is formed by $512 \times 512 \times 60/2$ fcc grid points.

Figures 12.17 and 12.18 show two axial slices of the VB dataset and corresponding slices of a 4-segmentation of it. Our segmentation program produced this 4-segmentation of the VB dataset containing 7,864,320 fcc grid points in 263 s or approximately 33 μs per spel.

One should observe that even though the attenuation coefficients of the reconstructed spels in the lung area are in the same range (for the current

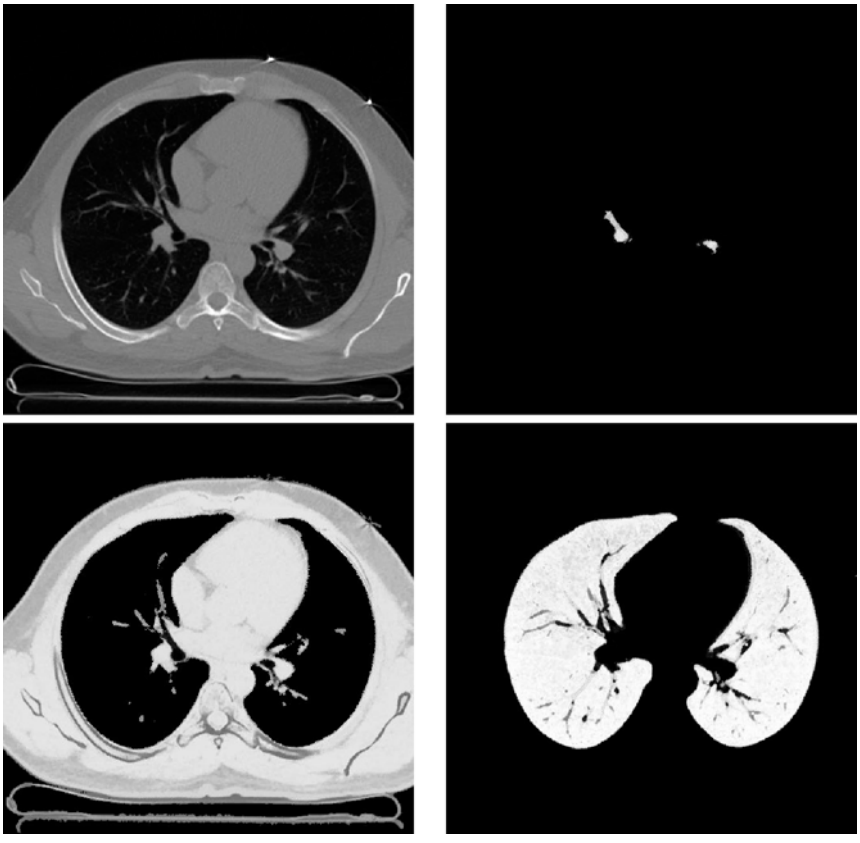


Figure 12.17: An axial slice from the VB dataset volume placed on the fcc grid and three maps of a 4-segmentation of it. (All four images were interpolated to the sc grid for display purposes.)

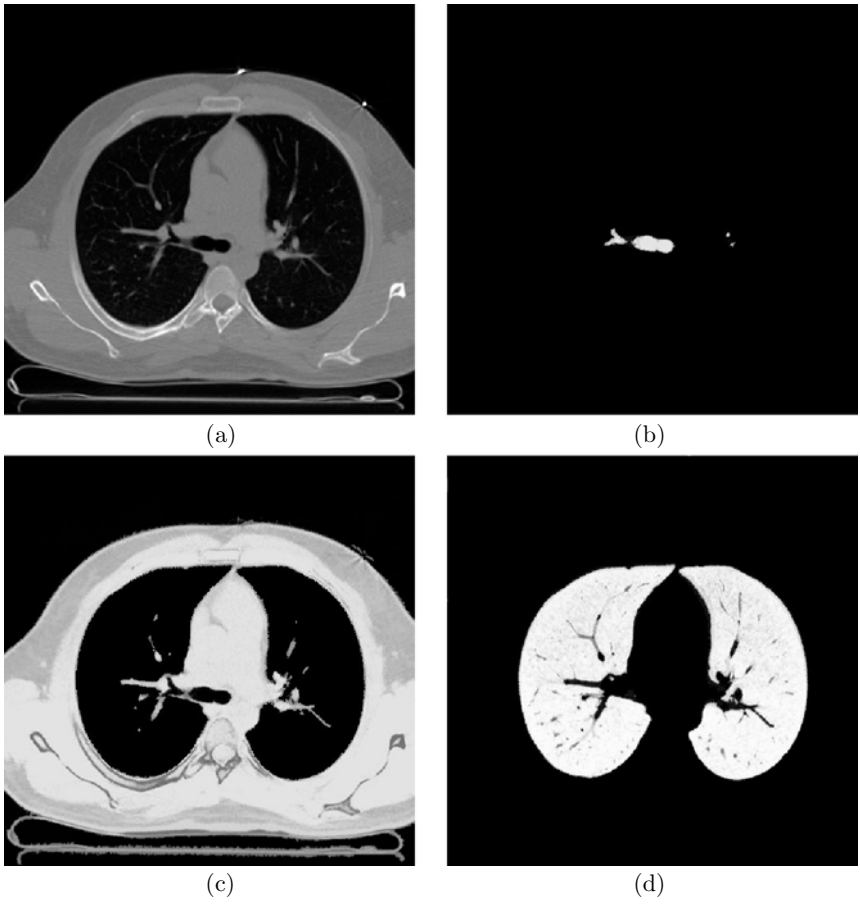


Figure 12.18: An axial slice from the VB dataset volume placed on the fcc grid and three maps of a 4-segmentation of it. (All four images were interpolated to the sc grid for display purposes.)

display window settings) as those in the bronchi and trachea, the placement of seeds in the areas of the lung close to bronchial junctions stop the leakage of the trachea–bronchi object (top right) into the lung object (bottom right). Figure 12.19 shows a 3-D view of the segmented trachea and bronchial tubes.

12.5 Conclusion

We have proposed a general, efficient, and easy to use semiautomatic algorithm for performing the simultaneous fuzzy segmentation of multiple objects. These

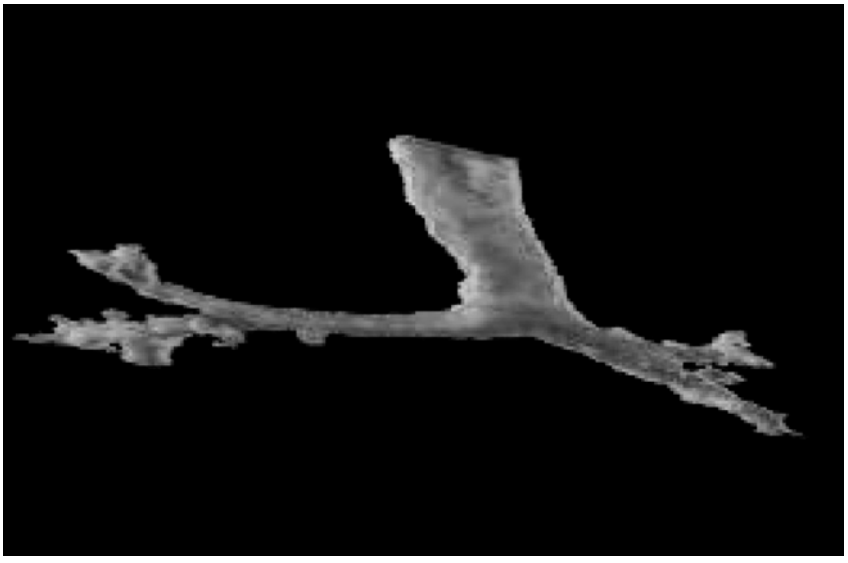


Figure 12.19: A 3-D view of the segmented trachea and bronchi shown in Figs. 12.17 and 12.18.

characteristics make the proposed method a valuable tool for interactive segmentation, since a low-quality segmentation (as judged by the user) can be corrected by the removal or introduction of new seed spels and a series of segmentations can be produced until a satisfactory one is achieved. The method can also be transformed into a fully automatic one if sufficient prior information is available pertaining the objects to be segmented.

12.6 Acknowledgements

We thank T. Yung Kong for his contributions to this work. This research has been supported by NIH grant HL70472 (GTH and BMC) and CAPES-BRAZIL (BMC).

Questions

1. *Characterize Algorithm 1 according to category and interactivity level.*

   Assuming that the unit of length is such that the distance between the nearest distinct points in the Vs of Fig. 12.20 is 1, we can define a fuzzy

The following figure and definitions are pertinent to Questions 2 to 6.

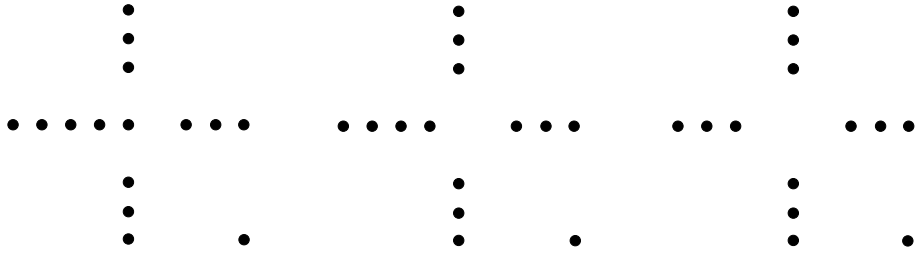


Figure 12.20: Three examples of a set of spels V; in each case the spels are dots in the plane and $V = T \cup L \cup R \cup B \cup \{o\}$, where T contains the top three dots, L contains the five (a), four (b), or three (c) horizontally centered dots on the left, R contains the three horizontally centered dots on the right, B contains the three vertically centered dots on the bottom, and o is the dot on the bottom-right.

spel affinity on V as any of the following:

$$\psi(c,d) = \begin{cases} 0 & \text{if } c = d, \\ 1/\|c-d\| & \text{otherwise}, \end{cases} \tag{12.21}$$

where $\|c-d\|$ is the Euclidean distance between the dots c and d,

$$\bar{\psi}(c,d) = \begin{cases} \psi(c,d) & \text{if } \|c-d\| \le 3, \\ 0 & \text{otherwise}, \end{cases} \tag{12.22}$$

and

$$\bar{\bar{\psi}}(c,d) = \begin{cases} 1/3 & \text{if } \|c-d\| \le 4, \\ 0 & \text{otherwise}. \end{cases} \tag{12.23}$$

2. *Are the sets V of Fig. 12.20 ψ-connected? If not, why?*

3. *Are the sets V of Fig. 12.20 $\bar{\psi}$-connected or $\bar{\bar{\psi}}$-connected? If not, why?*

4. *Consider the seeded 2-fuzzy graph $(V, \Psi, \mathcal{V})$ where V is the set (a) of Fig. 12.20, $\Psi = (\psi, \psi)$, V_1 contains the leftmost spel of V and V_2 contains the rightmost spel of V. Compute the 2-segmentation σ using Theorem 1.1.*

5. *Does the 2-segmentation σ change if we use $\Psi = (\psi, \bar{\bar{\psi}})$?*

6. *Is* $(V, (\psi, \bar{\psi}), \mathcal{V})$, *where* V_1 *contains the leftmost spel of* V *and* V_2 *contains the rightmost spel of* V, *a connectable* 2*-fuzzy graph for any of the sets of Fig. 12.20?*

7. *What does the concept of blocking of chains mean?*

8. *Why should one use the fcc grid instead of the traditional sc (cubic) grid?*

9. *Suppose that the fuzzy spel affinities defined for a specific application can only assume values from a small set (around* 1000 *elements). Discuss an alternative data structure for implementing the algorithm more efficiently.*

The following definitions are pertinent to Questions 10 and 11.
Using the notation of this chapter, the Relative Fuzzy Connectedness (RFC)[3] *of [27] defines a* 2*-segmentation as follows. For* $1 \leq m \leq 2$ *and for any* $c \in V$, *let* μ_m^c *denote the* ψ*-strength of the strongest chain from (the unique element of)* V_m *to* c. *Then, let*

$$\sigma_1^c = \begin{cases} \mu_1^c & \text{if } \mu_1^c > \mu_2^c, \\ 0 & \text{otherwise,} \end{cases} \tag{12.24}$$

$$\sigma_2^c = \begin{cases} \mu_2^c & \text{if } \mu_1^c \leq \mu_2^c, \\ 0 & \text{otherwise,} \end{cases} \tag{12.25}$$

and $\sigma_0^c = max\left\{\sigma_1^c, \sigma_2^c\right\}$ *for all* $c \in V$.

The Iterative Relative Fuzzy Connectedness (IRFC) of [27] produces a sequence ${}^0\psi_2, {}^1\psi_2, \ldots$ *of spel-adjacencies and a sequence of* ${}^0\sigma, {}^1\sigma, \ldots$ *of 2-segmentations defined as follows:* ${}^0\psi_2 = \psi$ *and* ${}^0\sigma$ *is the 2-segmentation defined by RFC. Now assume that, for some* $i > 0$, *we have already obtained* ${}^{i-1}\psi_2$ *and* ${}^{i-1}\sigma$. *For all* $c, d \in V$, *we define*

$${}^i\psi_2(c,d) = \begin{cases} 1 & \text{if } c = d \\ 0 & \text{if } {}^{i-1}\sigma_1^c > 0 \text{ or } {}^{i-1}\sigma_1^d > 0, \\ \psi(c,d) & \text{otherwise.} \end{cases} \tag{12.26}$$

[3]The definitions of RFC and IRFC of [27] are restricted to 2-fuzzy graphs where $\psi_1 = \psi_2$ with a single seed spel per object.

Then ${}^i\sigma$ *is defined just as* σ *is defined in RFC using (12.24) and (12.25), but with* μ_m^c *replaced by* ${}^i\mu_m^c$ *everywhere. Whenever* ${}^i\sigma = {}^{i-1}\sigma$*, then that 2-segmentation is considered to be the final output of IRFC.*

10. *Consider the seeded* 2*-fuzzy graph* $(V, \Psi, \mathcal{V})$ *where* V *is the set (c) of Fig. 12.20,* $\Psi = (\psi, \psi)$*,* V_1 *contains the leftmost spel of* V *and* V_2 *contains the rightmost spel of* V*. Compute the* 2*-segmentations* σ *using Theorem 1.1 and RFC and compare them.*

11. *Consider the seeded* 2*-fuzzy graph* $(V, \Psi, \mathcal{V})$ *where* V *is the set (a) of Fig. 12.20,* $\Psi = (\psi, \psi)$*,* V_1 *contains the leftmost spel of* V *and* V_2 *contains the bottommost spel of* B*. Compute the* 2*-segmentations* σ *using Theorem 1.1 and IRFC and compare them.*

Bibliography

[1] Canny, J. F., A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, pp. 679–698, 1986.

[2] Gonzalez, R. C. and Woods, R. E., Digital Image Processing, Addison-Wesley, Reading, MA, 1992.

[3] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Computer Vision, Vol. 1, pp. 321–331, 1988.

[4] Cohen, L. D., On active contour models and balloons, CVGIP: Image Understanding, Vol. 53, pp. 211–218, 1991.

[5] Geman, D., Geman, S., Graffigne, C., and Dong, P., Boundary detection by constrained optimization, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, pp. 609–628, 1990.

[6] Hofmann, T., Puzicha, J., and Buhmann, J. M., Unsupervised texture segmentation in a deterministic annealing framework, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 20, pp. 803–818, 1998.

[7] Mumford, D. and Shah, J., Optimal approximations by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math., Vol. 42, pp. 577–684, 1989.

[8] Ronfard, R., Region-based strategies for active contour models, Int. J. Comput. Vision, Vol. 13, pp. 1374–1387, 1994.

[9] Malladi, R., Sethian, J. A., and Vemuri, B. C., Shape modelling with front propagation: A level set approach, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 17, pp. 158–175, 1995.

[10] Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E., and Willsky, A., A shape-based approach to the segmentation of medical imagery using level sets, IEEE Trans. Med. Imag., Vol. 22, pp. 137–154, 2003.

[11] Udupa, J. K. and Samarasekera, S., Fuzzy connectedness and object definition: Theory, algorithms and applications in image

segmentation, Graph. Models Image Proc., Vol. 58, pp. 246–261, 1996.

[12] Carvalho, B. M., Gau, C. J., Herman, G. T., and Kong, T. Y., Algorithms for fuzzy segmentation, Pattern Anal. Appl., Vol. 2, pp. 73–81, 1999.

[13] Johnson, S. C., Hierarchical clustering schemes, Psychometrika, Vol. 32, pp. 241–254, 1967.

[14] Moghaddam, H. A. and Lerallut, J. F., Volume visualization of the heart using MRI 4D cardiac images, J. Comput. Inform. Tech., Vol. 6, pp. 215–228, 1998.

[15] Rice, B. L. and Udupa, J. K., Clutter-free volume rendering for magnetic resonance angiography using fuzzy connectedness, Int. J. Imag. Syst. Tech., Vol. 11, pp. 62–70, 2000.

[16] Saha, P. K., Udupa, J. K., and Odhner, D., Scale-based fuzzy connected image segmentation: Theory, algorithms and validation, Comput. Vision Image Understanding, Vol. 77, pp. 145–174, 2000.

[17] Udupa, J. K., Wei, L., Samarasekera, S., Miki, Y., van Buchem, M. A., and Grossman, R.I., Multiple sclerosis lesion quantification using fuzzy-connectedness principles, IEEE Trans. Med. Imag., Vol. 16, pp. 598–609, 1997.

[18] Rosenfeld, A., Fuzzy digital topology, Inform. Control, Vol. 40, pp. 76–87, 1979.

[19] Herman, G. T., Geometry of Digital Spaces, Birkhäuser, Boston, MA, 1998.

[20] Dellepiane, S. G., Fontana, F., and Vernazza, G. L., Nonlinear image labeling for multivalued segmentation, IEEE Trans. Image Process., Vol. 5, pp. 429–446, 1996.

[21] Ahuja, N., Dot pattern processing using Voronoi neighborhoods. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 3, pp. 336–343, 1982.

[22] Zahn, C. T., Graph-theoretic methods for detecting and describing Gestalt clusters, IEEE Trans. Comp., Vol. 1, pp. 68–86, 1971.

[23] Jain, A. K., Murty, M. N., and Flynn, P. J., Data clustering: A review, ACM Comput. Surveys, Vol. 31, pp. 264–323, 1999.

[24] Gower, J. C. and Ross, G. J. S., Minimum spanning trees and single linkage cluster analysis, Appl. Statist., Vol. 18, pp. 54–64, 1969.

[25] Pal, S. K. and Majumder, D. K. D., Fuzzy Mathematical Approach to Pattern Recognition, Wiley Eastern, L., New Delhi, India, 1986.

[26] Cormen, T. H., Leiserson, C. E., and Rivest, R. L., Introduction to Algorithms, MIT Press, Cambridge, MA, 1990.

[27] Udupa, J. K., Saha, P. K., Udupa, J. K., and Lotufo, R. A., Fuzzy connected object definition in images with respect to co-objects, In: Proc. SPIE, Bellingham, WA, Vol. 3661: Image Processing, Hanson, K. M., ed., pp. 236–245, 1999.

[28] Carvalho, B. M., Herman, G. T., and Kong, T. Y., Simultaneous fuzzy segmentation of multiple objects, In: Electronic Notes in Discrete Mathematics, Vol. 12, Del Lungo, A., Di Gesù, V., and Kuba, A., eds., Elsevier, Amsterdam, 2003. http://www.elsevier.com/gejng/31/29/24/71/23/59/endm12002.pdf.

[29] Herman, G. T. and Carvalho, B. M., Multiseeded segmentation using fuzzy conectedness, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 23, pp. 460–474, 2001.

[30] Garduño, E., Vizualization and Extraction of Structural Components from Reconstructed Volumes, Ph.D. Thesis, Bioengineering Program, University of Pennsylvania, 2002.

[31] Pollak, I., Willsky, A. S., and Krim, H., Image segmentation and edge enhancement with stabilized inverse diffusion equations, IEEE Trans. Image Proc., Vol. 9, pp. 256–266, 2000.

[32] Koepfler, G., Lopez, C., and Morel, J.-M., A multiscale algorithm for image segmentation by variational method, SIAM J. Numer. Anal., Vol. 31, pp. 282–299, 1994.

[33] Petersen, D. P. and Middleton, D., Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces, Inform. and Control, Vol. 5, pp. 279–323, 1962.

[34] Carvalho, B. M., Garduño, E., and Herman, G. T., Multiseeded fuzzy segmentation on the face centered cubic grid, In: Advances in Pattern Recognition: Second International Conference, ICAPR 2001, Rio de Janeiro, Brazil, 2001. LNCS Vol. 2013, Singh, S., Murshed, N., and Kropatsch, W., eds., Springer-Verlag, pp. 339–348, 2001.

[35] Carvalho, B. M., Cone-Beam Helical CT Virtual Endoscopy: Reconstruction, Segmentation and Automatic Navigation, Ph.D. Thesis, Computer and Information Science Program, University of Pennsylvania, 2003.

[36] IBM, Visualization Data Explorer User's Guide, Version 3 Release 1 Modification 4. http://www.opendx.org/support.html.

Chapter 13

Computer-Aided Diagnosis of Mammographic Calcification Clusters: Impact of Segmentation

Maria Kallergi,[1] *John J. Heine*,[1] *and Mugdha Tembey*[1]

13.1 Introduction

Medical image analysis is an area that has always attracted the interest of engineers and basic scientists. Research in the field has been intensified in the last 15–20 years. Significant work has been done and reported for breast cancer imaging with particular emphasis on mammography. The reasons for the impressive volume of work in this field include

(a) increased awareness and education of women on the issues of early breast cancer detection and mammography,

(b) the potential for significant improvements both in the fields of imaging and management, and

(c) the multidisciplinary aspects of the problems and the challenge presented to both engineers and basic scientists.

The importance of mammography and computer applications in mammography has been and continues to be the topic of numerous workshops, conferences, and publications [1, 2]. There seems to be sufficient evidence that mammography helps in the early detection of breast cancer although there are occasionally

[1]Department of Radiology, H. Lee Moffitt Cancer Center & Research Institute, University of South Florida, Tampa, FL 33612-4799

arguments that support an opposite view [3]. It should be noted that breast cancer was the second major cause of death for women in 2003 and mammography has been responsibly for a mortality reduction of 20–40% [4, 5]. Despite its success, mammography still has a false negative rate of 10–30% and great variability [6].

Calcifications are one of the main and earliest indicators of cancer in mammograms. They are present in 50–80% of all mammographically detected cancers but pathologic examinations reveal an even greater percentage [7]. Most of the minimal cancers and *in-situ* carcinomas are detected by the presence of calcifications [7]. A review of the literature on missed breast cancers indicates that calcifications are not commonly found among the missed lesions [8]. Although perception errors are not excluded, particularly in the case of microcalcifications (size < 1 mm), the technique of screen/film mammography (SFM) has been significantly improved over the years offering high-contrast and high-resolution mammograms that make calcification perception relatively easy. A greater and continuing problem for radiologists, with a major impact on the specificity of the diagnosis, is the mammographic differentiation between benign and malignant clustered calcifications. Almost all cases with calcifications are recommended for biopsy but only about 15–34% of these prove to be malignant [9]. The biopsies necessary to make the determination between benign and malignant disease represent the largest category of induced costs of mammography screening and a major source of concern for radiologists, surgeons, and patients. The advent of full field direct digital mammography will probably amplify this problem by providing more details and revealing breast abnormalities at very early stages [10].

In the last 20 years, researchers have developed various computer schemes for analyzing mammograms with calcifications, masses, and other breast abnormalities in an effort to improve mammography and breast cancer detection and diagnosis [11]. Computer algorithms can be divided in three groups depending on their final goal: detection, diagnosis, and prognosis methodologies [12]. The majority of the effort to-date has been focused on the development of detection tools, namely tools that point out to the primary reader suspicious areas associated with calcification clusters or masses that may warranty further review. The outcome of the intensive research on detection has led to three commercial, FDA approved systems for computer-aided detection (CADetection) of calcifications and masses; two more manufacturers were in the process of applying

for FDA approval as of this writing [2]. The commercial CADetection systems play the role of a virtual "second reader" by highlighting suspicious areas for further review and evaluation by the human observer [2].

Research on computer tools for diagnosis has been lacking behind but is now gaining momentum. The goal of a CADiagnosis system is to aid in the differentiation between benign and malignant lesions identified previously by a human observer [13] or a CADetection technique [14]. Such systems are not fully tested yet for clinical efficacy but are very promising and may provide significant aid to the mammographer in the form of a "second opinion."

Finally, computer-aided prognosis (CAP) tools appear in the horizon and are beginning to be explored as the next step in computer applications for breast imaging. Certainly, the variety of problems encountered in the detection, management, treatment, and follow-up of breast cancer patients leave several unexplored areas where computer applications could be clinically useful with major benefits to health care delivery and patient management.

Although the goals of automated detection and diagnosis are different, the actual detection and classification tasks are not always separate [15]. Almost all modern detection algorithms contain modules that discriminate true from false signals, calcification-like or mass-like artifacts from true calcifications or masses, isolated or single calcifications from clustered ones, even benign from malignant lesions in order to point only to malignant ones [16]. Most of the computer-based diagnosis techniques rely on the human observer to provide the detection step and/or the classification features [13, 17]. Few, however, incorporate segmentation and detection with pattern recognition processes in order to provide an automated, seamless approach that yields detection as well as likelihood of malignancy [18].

CADiagnosis methodologies that aim at the automatic differentiation of benign from malignant calcifications use a variety of mathematical descriptors that represent or correlate with one or more clinical findings, demographic information, or purely technical image characteristics. Reported algorithms usually employ combinations of morphological, texture, and intensity features as well as patient-related, demographic information [13, 14, 17, 19]. A valuable, comprehensive summary of reported techniques is given by Chan *et al.* [18]. Most of these methods are successful particularly when compared to the diagnostic performance of the human readers and their positive predictive value that is relatively low. Jiang *et al.* [14] have reported one of the first ROC

studies with a fully automatic classification method showing positive results. Despite advances in the field, several questions remain regarding the robustness of the current classification techniques, while their clinical benefits remain largely under-investigated.

This chapter looks into the types of CADiagnosis algorithms where segmentation, detection, and classification are combined in a seamless methodology that yields an outline of detected clustered calcifications and a likelihood of their malignancy. Furthermore, we look into algorithm designs where segmentation may play a critical role on the classification performance and we try to address issues related to the segmentation process and its validation. CADiagnosis methodologies that include a segmentation component and rely heavily on features that are extracted from the segmentation output could be very robust but entail a significant risk of being dependent on digitization conditions, i.e., laser vs. charge-coupled device (CCD)-based film scanners that vary in dynamic resolution characteristics, and on the source of digital data, i.e., digitized film vs. direct digital images. In addition, performance may depend on the nature of the segmented signals, i.e., false vs. true segmentations, artifacts vs. true objects, and even the criteria applied for the estimation of performance parameters including the type of gold standard available. This chapter reviews the work done by these investigators on CADiagnosis for mammography and breast microcalcification clusters, emphasizes segmentation issues, and reviews their impact on classification.

13.2 CADiagnosis Algorithm Design

The diagram in Fig. 13.1 shows the major components of a CADiagnosis algorithm aiming at the differentiation between benign and malignant lesions. Based on this diagram, one may distinguish two major pathways to algorithm development:

1. In one approach, a fully automated scheme is developed. Namely, the algorithm includes automated detection, feature selection, and classification modules. In this case, the diagnosis component of the algorithm may be considered as preceded by a CADetection component for an overall automated process.

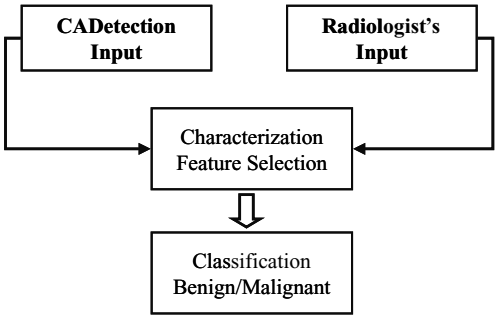


Figure 13.1: Block diagram of two possible CADiagnosis approaches. A combination of the two major pathways may also be used in CADiagnosis development.

2. In the second approach, automated classification is performed while detection of the lesion and classification features are provided by the human observer. In this case, all jobs prior to the classification step are "manually" done and observer-selected features are given as inputs to the classification module.

A CADiagnosis algorithm may be applied to various types of image data. Its design often depends on the specific application and data source. The diagram of Fig. 13.2 shows the various data types used to-date for CADetection and CADiagnosis applications. Most of the reported work has been focused on 2-D digitized film of a single breast view, either full size images or smaller regions of interest (ROIs) that contain only the lesion to be analyzed.

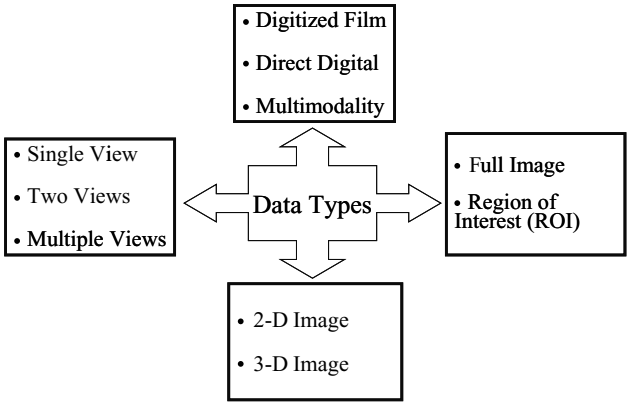


Figure 13.2: Diagram of potential data types used in CADiagnosis applications. Single types or combinations may be used.

In the literature, automated detection and diagnosis appear to be separated processes. However, as we discover when reviewing the elements of the various CADetection algorithms, more often than not, CADetection includes classification modules that allow the differentiation of true from false detected signals or benign from malignant signals in an effort to eliminate or reduce false positive signals and only point out to truly suspicious or cancerous-like areas. So, there is an inherent classification process in CADetection methods that is designed to remove signals not likely to be related to cancer. This process, however, is not very successful and has insufficient discriminatory power as indicated by the relatively large number of false positive (FP) signals appearing in the CADetection output (in both commercial and research systems) that are often related to benign conditions, e.g., benign calcifications and normal lymph nodes, and not only normal tissues. Another indication for the unsuccessful discrimination is the expressed frustration and concern among users of the CADetection systems for the large number of FP markers for either calcifications or masses that may confuse interpretation. Hence, the classification process included in CADetection algorithms for FP reduction may be considered as partial classification or not fully optimized as it does not provide accurate discrimination between benign and malignant lesions. In a CADiagnosis algorithm that follows the first approach above, CADetection is usually designed without a benign/cancer classification step, it aims at detecting all potential and true signals in an image (benign or malignant), and classification is performed at the final stage of the methodology with a dedicated component.

This chapter presents a CADiagnosis scheme that follows the first approach for the diagnosis of mammographic microcalcification clusters. The scheme was designed to reproduce a clinical visual analysis system that has shown significant success in the evaluation and diagnosis of calcifications clusters based on their morphology and distribution [20–22]. Hence, characterization and feature selection was confined to the morphological and distributional characteristics of calcifications excluding intensity or texture properties [18]. Another motivation for limiting feature selection to certain domains was the need to establish a validation tool for algorithms applied for the segmentation of calcifications and calcification clusters in mammograms that would avoid the path of "ground truth" comparisons. It is well known that there is significant ambiguity in the ground truth information provided by human observers. As a result, segmentation validation becomes highly uncertain. Our hypothesis was that calcification

segmentation methods are better evaluated indirectly based on a classifier's performance, if the classifier uses features defined only by the segmented objects.

13.2.1 Calcification Characteristics and Clinical Visual Analysis System

The clinical visual analysis system that formed the basis for the design of our algorithm and guided our feature selection is described in detail elsewhere [20–22]. It is based on several descriptors of the morphology and distribution of individual and clustered calcifications on mammograms. The number of calcifications in a cluster is not considered by itself a clear indicator of benign or malignant disease but when combined with other characteristics can increase or decrease suspiciousness [23–25]. The combination of all these properties by the human observer was shown to yield a sensitivity of 97.6% (correct identification of cancers associated with calcifications) and a specificity of 73.3% (correct identification of benign cases associated with calcifications) [20, 22].

The Breast Imaging Reporting and Data System (BIRADS) Lexicon of the American College of Radiology (ACR) was established in 1993 in an effort to standardize and improve mammographic interpretation. BIRADS was based on the clinical visual system of analysis. The recommended BIRADS descriptors for calcifications and calcification clusters are summarized in Table 13.1 [26]. Overall, there is strong evidence that morphology and distribution are two of the most important clinical aids in making the diagnosis of mammographic calcifications.

In clinical practice, a radiologist makes the final diagnosis of the detected calcifications based on the BIRADS characteristics, demographic information, and associated mammographic findings. However, inter- and intraobserver variability in the assignment of morphological features to the identified calcifications and ambiguity in the interpretation significantly degrades diagnostic performance. Hence, successful differentiation is limited among radiologists and can be as low as 20% leading to numerous unnecessary biopsies of cases with calcifications clusters [27].

Computer algorithms could translate and automate the clinical experience and thus assist the radiologist in this diagnostic task. An algorithm that provides information on the morphology, e.g., segments calcifications while preserving size and shape, and gives a likelihood of malignancy for a detected calcification cluster could be extremely valuable in mammogram interpretation and patient

Table 13.1: BIRADS descriptors for calcifications with associated genesis type [26]

Morphology or character	Skin (lucent centered)	B
	Vascular (linear tubular with parallel tracks)	B
	Coarse or popcorn like	B
	Large rod-like	B
	Round (larger than 0.5 mm)	B
	Eggshell or rim (thin walled lucent centered, cystic)	B
	Milk of calcium (varying appearance in projections)	B
	Dystrophic (irregular in shape, over 0.5 mm, lucent centered)	B
	Punctate (round smaller than 0.5 mm)	B
	Suture (linear or tubular, with knots)	B
	Spherical or lucent center (smooth and round or oval)	B
	Amorphous or indistinct	U
	Pleomorphic or heterogeneous granular	M
	Fine linear	M
	Fine linear branching	M
Distribution	Clustered	U
	Segmental	U/M
	Regional	U
	Diffuse/Scattered	B
	Linear	M
Number	1–5	U
	5–10	U
	>10	U

B = probably benign; M = suggestive of malignancy; U = uncertain.

management. To be clinically useful, the algorithm should perform real time, be robust, and have consistent performance at least comparable to the clinical visual analysis system [20, 22]. The algorithm described here was designed to meet the above requirements and two additional conditions: (a) The desired classification performance had to be achieved with the smallest possible set of features. (b) Feature selection would be initially limited to the morphological, distributional, and demographic domains; expansion to other domains would be considered only if performance did not reach desirable levels. The specific components of this scheme are shown in Fig. 13.3. The algorithm was implemented and tested on simulated calcification clusters, large sets of mammographic calcifications, and datasets of various image resolutions [20, 28, 29]. All studies demonstrated that the development of a classifier on morphological characteristics alone is a viable and reliable approach. They also supported our hypothesis

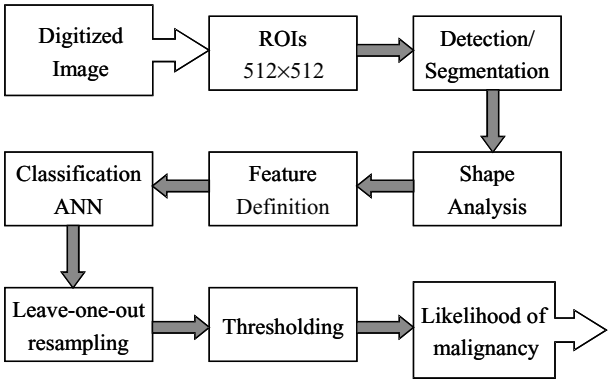


Figure 13.3: Flowchart of the CADiagnosis algorithm developed for the differentiation of benign from malignant microcalcification clusters in digitized, screen/film mammography [20].

that a classifier could be used as an indirect measure of segmentation performance [29]. The segmentation of the individual calcifications and the clusters with shape and distribution preservation was a critical stage in our methodology. Hence, in the following section, we will discuss the detection/segmentation stage in more detail with particular emphasis on the role of wavelets in this process.

13.3 Detection/Segmentation Stage

The terms segmentation and detection may be confusing for the reader not so familiar with the medical imaging vernacular. In some instances these terms may be used interchangeably, but other times not. We might consider segmentation as being a more refined or specialized type of detection. For instance, we may gate a receiver for some time increment and make a decision as to whether or not a signal of interest was present within the total time duration, but not care about exactly where the signal is within the time window; this may be defined as a detection task with a binary output of yes or no. Segmentation takes this a step further. With respect to the image processing, the detection task makes a decision if the abnormality is present, which in this case is a calcification. If, in addition, the detection provides some reasonable estimate as to the spatial location and extent of the abnormality, then we would say that the calcification

has been segmented. Thus, the segmentation process in mammography often results in a binary-labeled image with the probable calcification areas marked.

Before getting into the details of the techniques we implemented for the detection/segmentation stage of our CADiagnosis algorithm, a brief discussion of related bibliography is in order for the novice or beginner in the field to allow for a heads-up on study material of the suitable level. The list that follows is in no-way complete, or totally contemporary, but is comprised of useful citations (textbooks generally) that we have used extensively in our research and algorithm development.

Tolstov [30], Bracewell [31], and Brigham [32] are excellent sources for studying Fourier series and Fourier transforms—a prerequisite to understanding the wavelet transforms used in our approach. In particular, Brigham [32] provides a comprehensible treatment of the relations with the continuous Fourier transform, the discrete Fourier transform, and sampling theory. Similarly, Bracewell [33] gives a well-balanced treatment of standard imaging processing techniques.

Noise and filtering will be discussed in the following sections. Generally, the study of noise processes comes under many subject headings such as stochastic analysis, random signal analysis, or probability analysis. Again there are many diverse resources in this area and several provide many useful examples of random variable transformations and Fourier analysis of random signals [34–37]. An excellent treatment of transforms and probability analysis applications is given by Giffin [38].

Wavelet analysis may be looked at from a simple filtering approach as well as from an elegant mathematical framework that involves understanding multiresolution functional spaces. Again, there is massive published work in this area. Strang and Nguyen [39], Akansu and Haddad [40], and Vetterli and Kovacevic [41] are excellent sources for understanding wavelets from a filtering approach, which also include the multiresolution framework. The seminal work in wavelet theory may be found in the more mathematically sophisticated work of Daubechies [42].

Finally, in the sections below, we will discuss how mammograms are associated with power spectra that obey an inverse power law. This characteristic is associated with self-similarity, fractals, and chaos. We are not aware of any traditional textbooks that address power laws specifically but the work of Peitgen *et al.* [43], Wornell [44], and Turner *et al.* [45] may be useful; Peitgen *et al.* [43] cover many types of phenomena, while Wornell [44] and Turner *et al.* [45] are

specific to wavelet-based signal processing and 2-D image analysis, respectively. Note that the idea of self-similarity implies that things or events are invariant under a scale change. Wavelets have this property. Thus, it would seem natural to study self-similar noise fields (such as mammograms) with a self-similar analyzing transform (wavelets).

13.3.1 Wavelet Filtering

The term image enhancement is also a very general term, which encompasses many techniques. We have recognized that it is often used to describe the outcome of filtering. If the definition of enhancement is applying a process to the image that results in a "better" overall image appearance, then the term is a misnomer. Linear filtering blocks a portion of the true signal in most applications, which is probably not best defined as enhancement. In this section we provide a qualitative description of filtering. The only assumption we make here is that the reader understands Fourier analysis in one dimension. If so, the extension to two dimensions will be easily accomplished.

Things that change rapidly in the signal domain give rise to larger amplitudes for the sine waves that wiggle more quickly (high frequencies) in the Fourier expansion of the signal. Likewise, things that have long-range structure in the signal domain, give rise to larger amplitudes for the sine waves that wiggle slowly (low frequencies) in the Fourier expansion. Of course, there are many structures that lie in the middle of these extremes. The reader should keep in mind that the above descriptors are relative terms. Signals that are delta-function like will give rise to Fourier components across the entire spectrum. There is more to the story, because we are working in two dimensions. Let's consider a two-dimensional function or contrived image that is nothing more than an infinite vertical line (y-direction) of a few pixels in width in the other direction (x-direction) embedded in an empty 2-D field. Note that in the vertical direction there is no variation along the line indicating it will look as a low frequency signal in this direction (this may be considered as a sine wave with an infinite period, which may be considered as a DC component). If we approach this line from the horizontal direction, it appears as an abrupt change, for an instance, then it is flat for a few pixels, then another abrupt change takes place, and it is gone; that is, any horizontal slice will look like the rectangle function that is used in many Fourier analysis textbooks for examples. It takes two frequency coordinates to describe

this signal or any 2-D signal (image). Specifically, this signal is purely a DC signal in the y-direction when considering its Fourier composition and a sinc-type function in the other direction. Consequently, the transform is a sinc function along the f_x coordinate axis and about zero elsewhere; this may be deduced by considering the (f_x,f_y) coordinates and noting that the f_x component is zero everywhere except at $f_y = 0$. The following may be observed: (1) Linear structures in the vertical direction are likely to give rise to Fourier signatures in the f_x direction. The narrower the width the more spread out the contribution is in the Fourier f_x direction and the wider in width, the more contracted along the f_x direction. (2) Linear structures in the horizontal direction are likely to have significant Fourier signatures in the f_y direction and less in the f_x direction. (3) Taking this a step further, spots give rise to components in both coordinate directions. These examples are idealizations that may inspire the newcomer to Fourier analysis to observe what exactly the Fourier Transform is telling the user.

Filtering can be applied to set the stage for detection or segmentation. The basic idea is that there is some structure that we define as the signal of interest, which in our case is the localized calcified areas in mammograms termed "calcifications." These signals are surrounded (or embedded in) by other signals (in this case normal breast tissue) that may interfere with the ability to automatically detect them. In the best scenario, the signal of interest will have a frequency signature that is somewhat different than the background. If this is the case, filtering the image will pass the signal of interest (perhaps not intact) and block a portion of the background tissue. If this is successful, the filtered image will show a relatively more pronounced calcified area and a somewhat subdued background when compared with the raw image.

A simple somewhat contrived example of the usefulness behind filtering is proved here. Suppose we have a white 2-D $(n \times n)$ noise field with variance σ^2 and filter it with a perfect band pass filter. Can we say anything about the resulting noise power? The answer is yes. White noise by definition is a flat power spectrum (more correctly a constant power spectral density). For illustration, we will apply a perfect half-band filter to this field and calculate the resulting noise power. In the Fourier domain, the half band filter looks like a square box centered about zero of unit height with its sidewalls intersecting the frequency coordinates and the midway point. Fourier components within the box are passed when filtering and everything outside is blocked. Thus the total area in the Fourier domain is n^2, the pass-band area is $(n/2)^2$, and the blocked portion of

the Fourier domain is $n^2 - (n/2)^2 = \frac{3}{4}n^2$. Considering the transform properties, the resulting noise power is $\frac{3}{4}\sigma^2$. The important point here is that if the signal of interest has strong signatures in the lower part of the frequency spectrum, they will be passed almost intact while the noise would be heavily damped. This is an idealized situation that helps to understand the reasons for filtering.

In the following, a very brief description of wavelet analysis is presented. There is no way that we can give proper justice to this elegant theory of signal decomposition. Beware it took great minds many years to put wavelet analysis on such a beautiful foundation, which is now discussed as commonplace.

When considering the actual wavelet application, the wavelets are not expressed explicitly in an analytical form, but are expressed as two filter kernels corresponding to a weighted differencing operator and a weighted smoothing operator, which are complementary operators. In the literature these are referred to as the mother wavelet and associated scaling function. The forward transform is applied by alternative applications of the two kernels with down sampling interleaved between the applications. This procedure generates the wavelet expansion coefficients. The inverse transform is also achieved by repeated convolutions with two related filter kernels with up-sampling interleaved between the convolutions. The filtering aspect of the analysis is implemented by applying the forward transform and setting the desired coefficients to zero before applying the inverse transform.

The procedure described above may also be presented using the equivalent terms of dilation (or contraction) and translation of the mother wavelet function. For a given wavelet basis, there is really only one wavelet function or mother wavelet. The entire basis is constructed from translations and dilations of this wavelet. Spreading it out reduces the resolution and translating provides spatial information. The translations and dilations are not arbitrary, but are picked in a certain way from an orthogonal basis at multiple resolutions.

A way to view this is that the wavelet coefficients are really correlation figures of merit indicating how well the signal and a given region correlates with the particular version of the mother wavelet. The important thing to note here is that when the wavelet is spread out, the inner product with the signal at a given region (or spatial location) encompasses the length of the dilated wavelet. This implies that the wavelets coefficient holds information about the entire spatial region. As the wavelet spreads out, the frequency-band narrows. Thus, the analysis is better localized in frequency but worse localized in space.

The reverse argument applies when the wavelet is most contracted implying better spatial location but spread out in frequency. These ideas are fundamental to understanding both Fourier and wavelet analysis. For the purpose of this discussion, a very simple wavelet interpretation was developed and presented in the following paragraphs.

The wavelet expansion may be considered as a band pass filter network. The intact signal (raw image) is put into the mill and out come many filtered versions of the image. The orthogonal wavelet gives an expansion of the form:

$$F_0 = d_1 + d_2 + d_3 + d_4 + \cdots + d_j + F_j \tag{13.1}$$

where F_0 is the raw image, the d_js represent band pass versions of the raw image, and F_j is a blurred version of the raw image, which contains the DC and slow varying image attributes. These images are linearly independent, which amounts to perfect reconstruction and is one of the great strengths of wavelet analysis compared with just any band pass filter network. Each image of these expansion images may be divided further into three complimentary components expressed as vertical, horizontal, and diagonal components, which are also not correlated. Roughly speaking, the ds represent an octave sectioning (or fine to coarse image representation) of the frequency domain information. This can be observed by taking the Fourier transform of each expansion component individually and noting where each has appreciable Fourier amplitudes. Figure 13.4 shows the idealized division of the Fourier domain relative to the image expansion images.

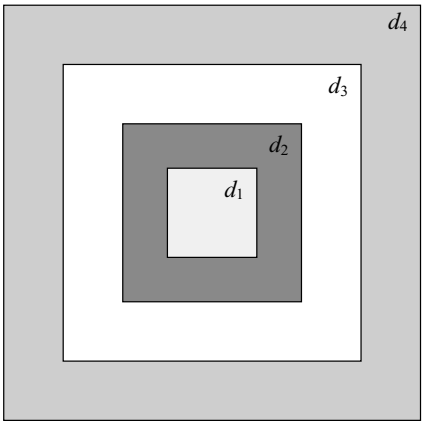


Figure 13.4: Idealized graphical representation of the first four band pass splittings of the raw image.

There are many orthogonal wavelet coefficients to choose from, and the band pass nature is not generally sharp indicating that the expansion components will have some shared frequency attributes; these cancel when performing the addition. Here is a simple rule of thumb: the shorter the wavelet filter kernel the less sharp the cutoffs are in the Fourier domain and the longer the support length the sharper the cutoffs. The strength of the two-channel or quadrature wavelet filter is the orthogonality of the expansion images. The price paid for orthogonality is the fixed-way the associated information is divided (octave sectioning).

13.3.2 Symmlet Wavelet Approach

As stated, there are many wavelet bases that can accomplish the detection and segmentation of calcified areas in mammogram and set the stage for the classification task. In a first application, we used a nearly symmetric wavelet. Our choice was guided by the close similarity between the wavelet profile and the calcification profiles (recall the correlation idea discussed above). First, the image is expanded as in Eq. (13.1). Deciding which components to discard is the crucial decision with this approach. This choice is dependent upon the calcification size (in both pixel width and actual linear measure) and the digital resolution. The term size is used in the average or expected sense. From the clinical view of a suspicious abnormality, calcifications of up to 1/2 mm are important. This translates into about 8–16 pixels in an image generated with a 35 μm/pixel digital resolution; our original work was performed at this high resolution and this experience will be discussed here [46]. In pilot studies, the d_3 and d_4 images demonstrated the largest calcification signatures relative to the background and were empirically selected for the process. Two pathways could be followed: (1) Add the two relevant expansion images together and perform the detection or (2) perform the detection in each image and combine the results afterwards. The latter option gave better sensitivity performance, since it gives the opportunity to detect some calcified areas twice. Specifically, small calcifications had a stronger signature in the d_3 images and large calcifications had stronger signatures in the d_4 images. Many calcifications had of course signatures spread across the two components.

Rather than impose a detection or decision rule on the process, we decided early on to see whether a parametric approach to decision making could be followed. Our pilot studies showed that the wavelet expansion images could be

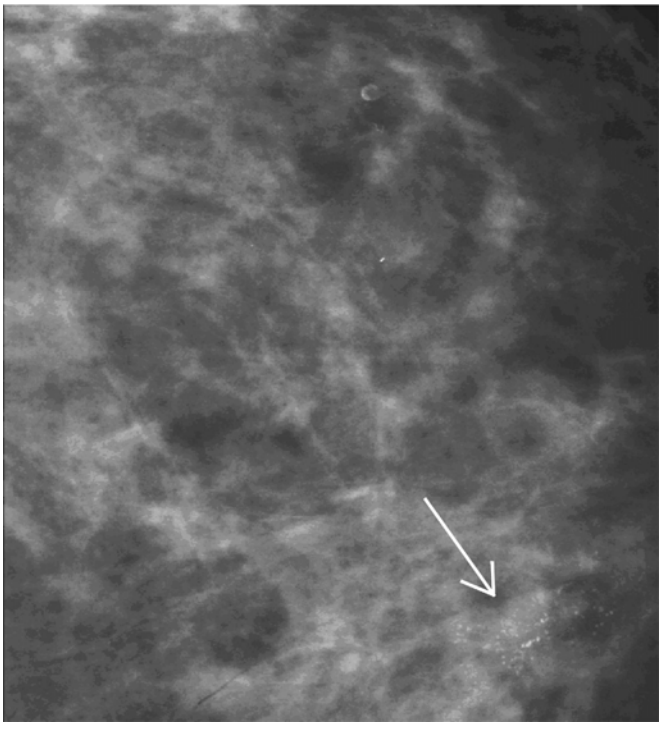


Figure 13.5: Representative mammographic section (2048 × 2048 pixels) with a malignant calcification cluster indicated by arrow. Image resolution is 35 μm and 12 bits/pixel.

approximated with parametric methods to characterize their empirical pixel distributions [46–48]. Calcifications, and calcification clusters, occupy a relatively small portion of the image when present; Fig. 13.5 shows a typical calcification cluster associated with cancer. Given the small area properties, the wavelet modeling produced essentially the parametric probability distribution functions (PDF) for normal tissue at multiple resolutions.

Theoretically, knowledge of the surrounding tissue PDF allows for the development of a statistical criterion to test against it using maximum likelihood analysis [49]. Our work indicated that the PDFs may be approximated from a family of parametric PDFs indexed by N; when $N = 1$, the PDF is Laplacian, and when N is large it tends to a normal distribution and the PDF is symmetric about the origin (zero mean).

We will not delve into this area of statistics here, but will indicate the approximations used. In this application, we ignored the pixel correlation within

a given expansion image and used a low order N approximation. Namely, if N is in the neighborhood of 3, the $N = 1$ approximation was applied to simplify the techniques. Likewise, before applying the maximum likelihood analysis, the expansion images where transformed to all positive values by taking the absolute value. Figure 13.6 shows the first three wavelet expansion images for the mammogram of Fig. 13.5.

Knowing the form of the PDF allows for the development of a test statistic, better described as a *summary statistic*, which follows from the maximum likelihood approach. However, the technique does not indicate how exactly to apply the test to the problem at hand. The maximum likelihood analysis indicated that the average was the test statistic. Tailoring this to our problem translated into sliding a small search window across the image matched in size to the expected calcification size. At each location the average was calculated, and if the local average deviated from the expected overall normal tissue average, it was labeled as suspicious and marked. Otherwise, the local region was set to zero. Thus, by systematically analyzing each local image region, most of the image was discarded and the potential abnormalities were labeled. A different search size window was used for the different images: an 8×8 pixel window was applied to the d_3 image and a 16×16 pixel window was applied to the d_4 image. The detection result yielded two binary images that for the most part were zero but were equal to 1 in areas corresponding to calcifications. The union of the two detections formed the initial total binary detection output. But detecting isolated calcifications was not the end of the process. Calcifications had to be grouped into clusters for further analysis. The clinical rule was followed here for grouping calcifications. Namely, a cluster was defined as three or more calcifications within a 1 cm^2 area [50]. Thus, in a second run, a larger search box was scanned across the binary-labeled detection image and isolated spots were set to zero.

The threshold(s) that give the best trade-off between labeling a normal area as suspicious (FP detection) and labeling a true calcified area as normal (false negative (FN) detection) must be probed with experimental methods. Generally, this requires a sample set of images with known calcification clusters and another sample set of images with no abnormalities at all. This image assemblage is processed repeatedly while varying the thresholds and calculating the performance rates: (1) the number of correctly identified calcification clusters (true positive (TP) detections) and (2) the number of FP clusters. In our work,

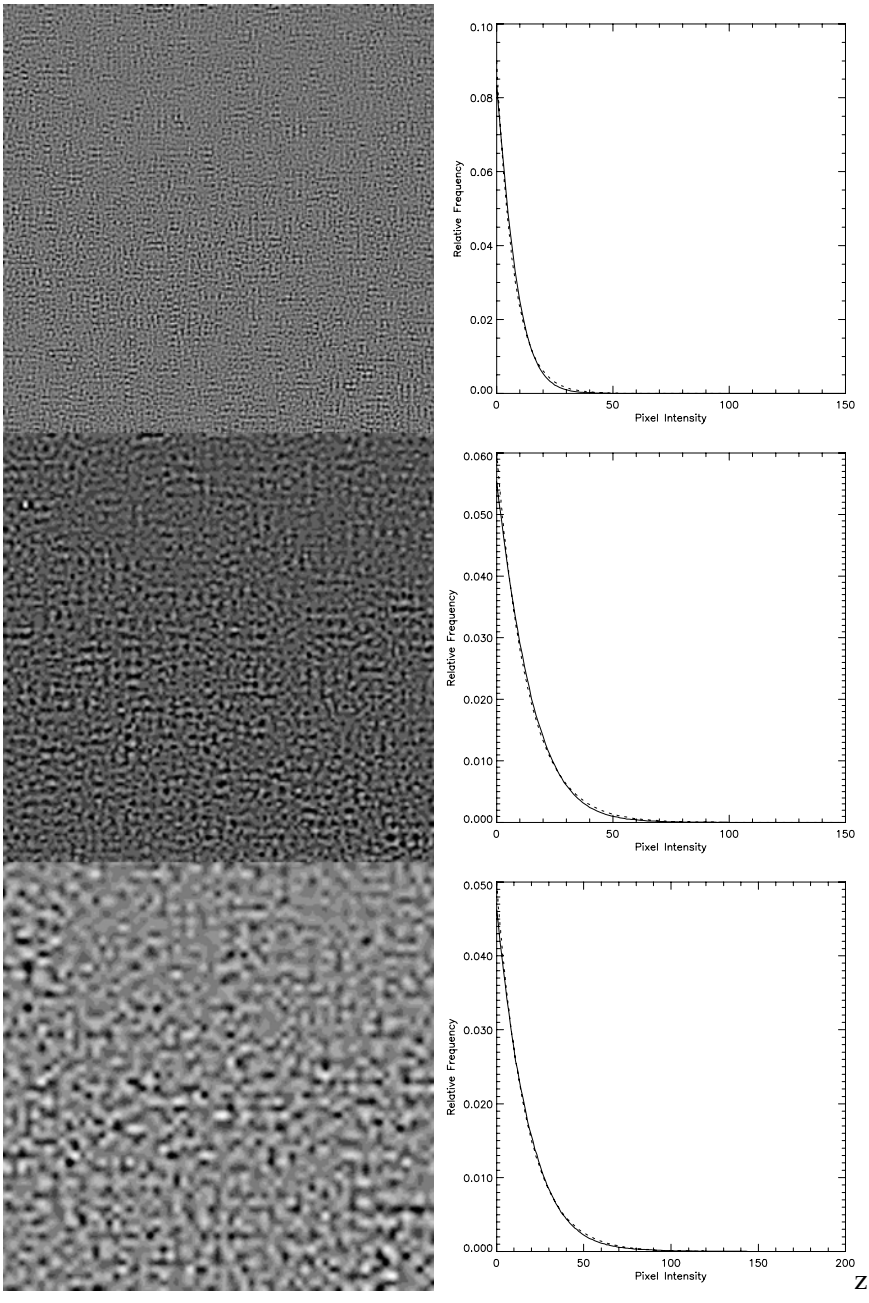


Figure 13.6: First three wavelet expansion images (d_1, d_2, and d_3 from top to bottom) corresponding to the raw image of Fig. 13.5. The probability modeling and empirical histograms are displayed on the right after taking the absolute value of the data. The theoretical curves are represented by solid lines and the empirical data by dashed lines. For viewing purposes, the images are 256×256 pixel sections cut form the original image of Fig. 13.5 but the probability modeling is derived form the entire image.

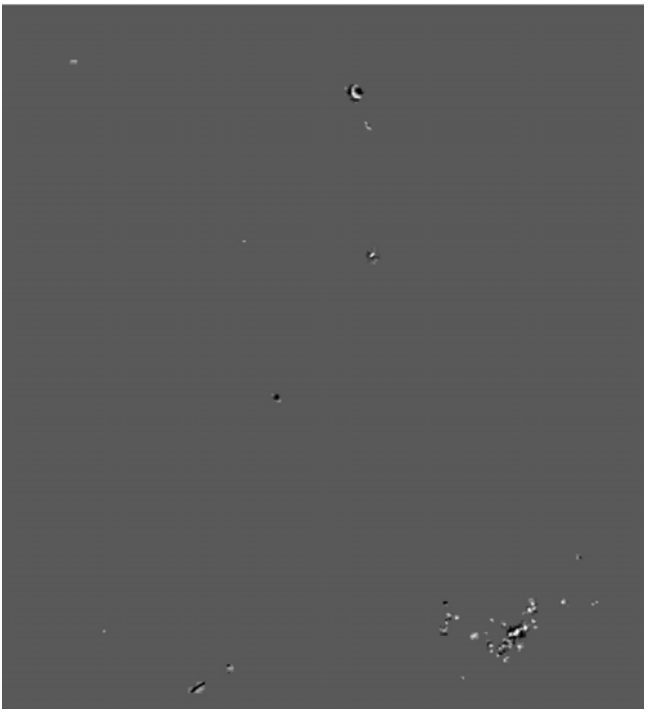


Figure 13.7: Detection output from the dual wavelet expansion image approach of Fig. 13.5. The binary mask has been projected into the sum of the first five d_j images, a process that gives better detail for further processing.

there were two thresholds associated with each detection stage that were varied independently. Figure 13.7 shows the output of this process for the cluster shown in Fig. 13.5.

As we shall see below, the classification algorithm we developed requires the analysis of calcification attributes that were not fully present in the binary detection representation of Fig. 13.7. Two options could lead to the desired representation: (a) The binary detection output may be used as a mask that points back to the calcification location in the raw image, or, more generally, to any other data representation. For example, classification analysis could be done on any combination of the d images in Eq. (13.1). (b) Perform an additional segmentation operation on the binary detection output that would extract the shape and distribution of the detected cluster(s) and allow their shape analysis necessary for the classification step. The second option was selected in this application and calcifications were segmented in the detection

image of Fig. 13.7 by applying an adaptive threshold process that is described in section 13.3.4.

It should be noted that the original symmlet wavelet method was developed and optimized for 35 μm and 12 bits/pixel image resolution. The algorithm was modified to be applicable to images of 60 μm and 16 bits/pixel, a resolution that was identified in separate experiments as optimum for morphology-based classification [28, 29]. Specifically, image resolutions of 60–90 μm/pixel were found to maintain the calcification shape and size characteristics on which to base feature selection for classification. Higher spatial resolutions, i.e. 30 or 35 μm, did not improve classification results but significantly increased computational intensity and image manipulation. Lower spatial resolutions, i.e., equal to or greater than 90 μm/pixel, degraded classification performance because of the losses in morphology and distribution of the detected calcifications. So, in the following section, we will discuss the limitations of the symmlet wavelet method for applications other than 35 μm and what led us to the design of a new filter for calcification detection independent of image resolution that significantly improved classification performance and robustness of the results.

13.3.3 The "Donut" Filter

For the moment, lets assume that the octave sectioning described previously divided the information in the best possible manner for detecting calcifications at a 35 μm image resolution. But what happens if we change resolution? It may be safe to say that, if the resolution is doubled or halved, we could pick different expansion images and our technique would still be "the best" within this hypothetical situation; that is, if the images were scaled to down 70 μm (lower resolution or halved), the d_1 and d_2 images would most likely be the most relevant choices. However, if the scaling was not applied by a factor of two (or half) we might end up with a less then optimal representation for detection purposes.

With this in mind and the outcome of the resolution studies mentioned earlier, we developed another band pass filter that we will refer to as the "donut" filter [51]. This filter has three infinitely adjustable parameters that control: (1) the central band location, (2) the band pass width, and (3) the sharpness of the cut-off. This filter is easily expressed in radial frequency coordinates as

$$\exp - \left[\frac{(f - m)}{\sqrt{2}\sigma} \right]^{l} \tag{13.2}$$

where f is a radial frequency variable, m locates the central band, σ controls the bandwidth, and l alters the cut-off. This filter has more adjustment leeway or freedom compared with the symmlet wavelet but at the cost of orthogonality. Certainly m and σ may be varied and a series of band pass filtered images may be generated the sum of which will not generate the raw image back. However, orthogonality is not always important.

In order to apply the new donut filter for calcification processing at an arbitrary image resolution, its operating parameters must be selected. There may be theoretical methods to approach this problem. However, an empirical method was used here based on signal-to-noise ratio analysis. For this, known calcification areas were hand-marked using a wide range of specimens from many images. Likewise, several normal image regions were labeled on the same images. The area markings for both normal and abnormal (calcification containing) tissue types were of the same size that changed depending on the particular image resolution. The calcifications were considered as "the signal" and the normal background tissue was considered as "the noise." These images were processed repeatedly while varying m and σ of Eq. (13.2) and calculating the power for the respective image regions. By averaging the power from the signal and noise regions across all images and forming the average signal-to-noise ratio provided the means for finding the best average or overall operating parameters. The parameter l was determined prior to this by a similar rationale and was set to 4; if $l = 2$, the profile was Gaussian. Figure 13.8 shows slices through the donut filter for $l = 2$ and $l = 4$. Note the difference in the cut-off properties.

Figure 13.9 presents a section of a mammogram digitized at 60 μm and 16 bits/pixel containing a calcification cluster associated with cancer. Figure 13.10 presents the symmlet wavelet output and Fig. 13.11 shows the donut filter output for the section of Fig. 13.9. The comparison of the two images in Figs. 13.10 and 13.11 indicates that more information was maintained with the donut filter than the symmlet wavelet approach something that impacted shape analysis and classification results as we shall see in the following sections.

In addition to the improved results obtained from the digitized mammograms, the donut filter could also be applied to full field digital mammography (FFDM) images with calcifications acquired with the new General Electric Senographe 2000D FFDM system. The result of the donut filter for a 100 μm FFDM image with calcifications is shown in the insert of Fig. 13.12. A prewhitening technique was applied first to the data before the application of the donut filter. The

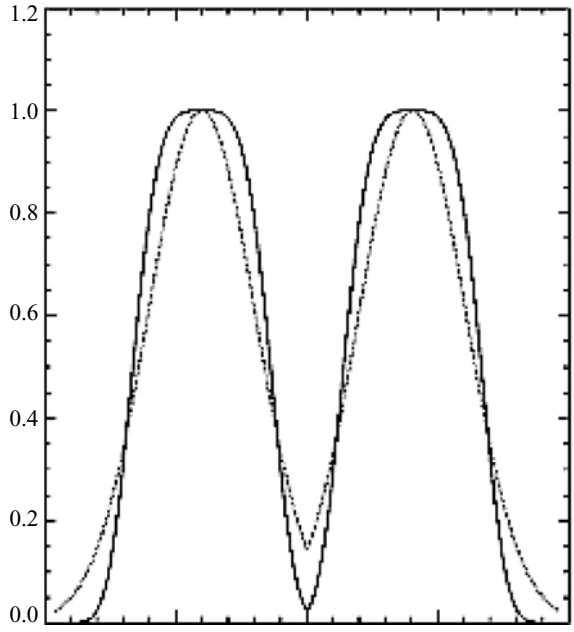


Figure 13.8: Slices through the donut filter in the Fourier domain for $l = 2$ (dashed curve) and $l = 4$ (solid curve) versions. Note the difference in the cut-off behavior of the two versions.

Figure 13.9: Original ROI (512×512 pixels) with a calcification cluster associated with cancer. The ROI was obtained from a screen/film mammogram digitized at 60 μm and 16 bits/pixel.

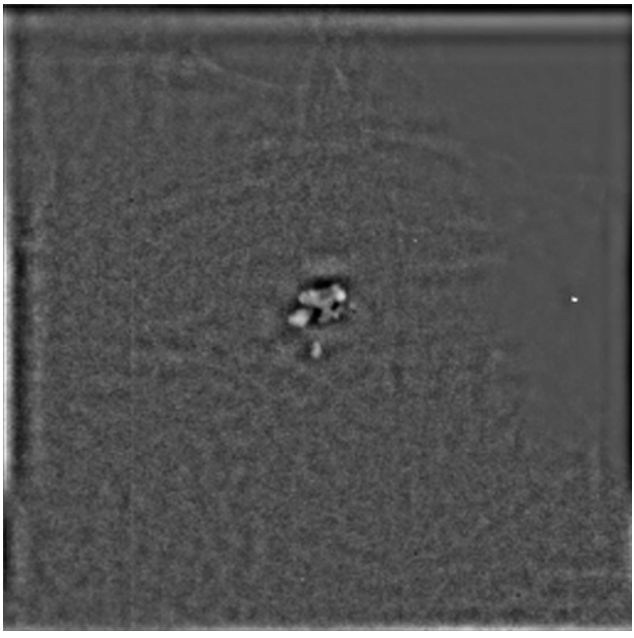


Figure 13.10: Output of the symmlet wavelet filter for the ROI of Fig. 13.9. Strong edge effects are present with this filter that often remain in the final segmentation step (see section 13.3.4) and interfere with classification.

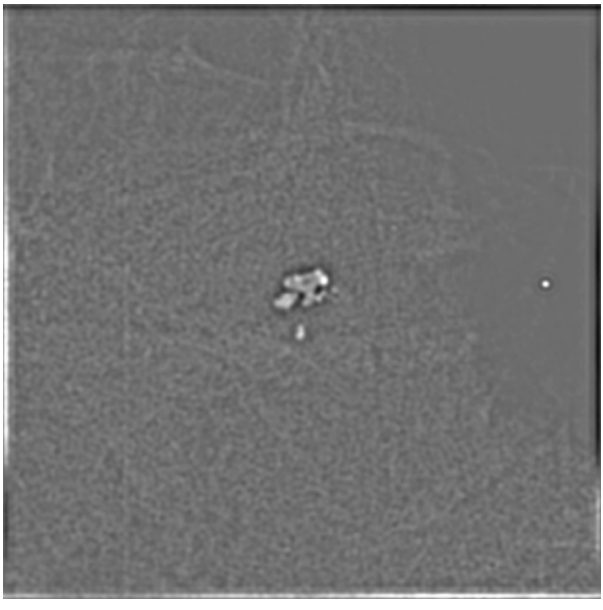


Figure 13.11: Output of the donut filter for the ROI of Fig. 13.9. Smaller edge effects are observed in this case and improved edge definition of the objects of interest.

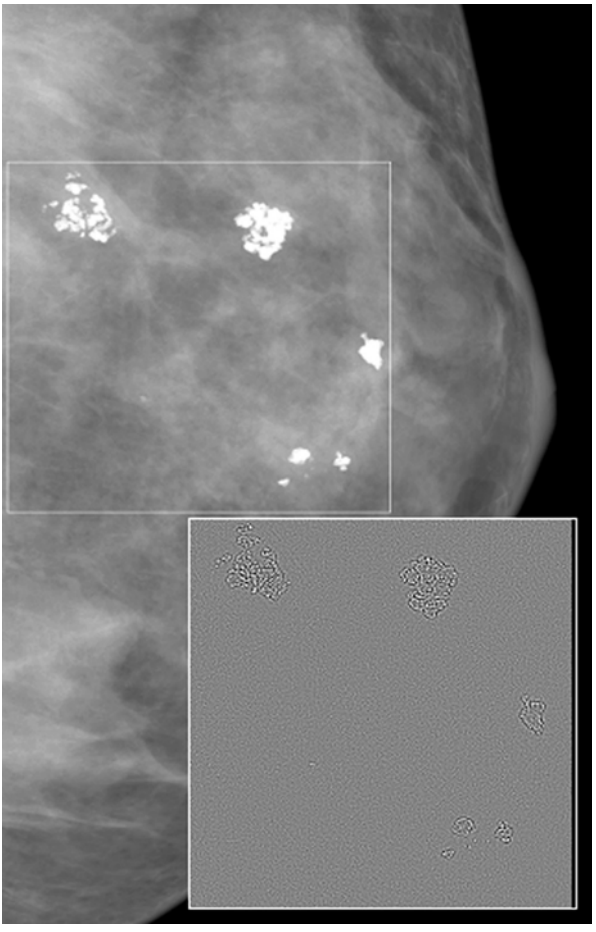


Figure 13.12: Main section of an FFDM image of the left breast of a patient with benign clustered calcifications enclosed in a white box. Image was acquired with GE's Senographe 2000D digital system at a resolution of 100 μm and 16 bits/pixel. The insert shows the region with the calcifications after filtering with the new donut filter in combination with a prewhitening approach. Note that the background is subdued (gray value information is removed) and calcifications remain as outlines that can be easily extracted.

prewhitening process amounts to removing the influence of the mammogram's natural spectral form before applying the filter. Preliminary results suggest that this preprocessing step could increase overall detection performance.

The edge artifacts present in both the wavelet-filtered and the donut-filtered image are due to the wrap around effect in the convolution process or the

periodic wrap around present in the discrete Fourier Transform. Basically, the filter kernel slides off one side of the image and appears on the other. Thus, for all practical purposes, the kernel slides over what appears as a discontinuity. The artifact appears more pronounced in the wavelet-filtered than the donut-filtered image, which may be due to the iterative convolution inherent to its application.

In all fairness, we have not discussed the characteristics of the actual mammograms, digitized or direct digital. Following, we will give a short description that will assist the reader in understanding the difficulties in processing mammograms either automatically (computer vision) or manually (human vision).

Evidence indicates that mammograms, regardless of resolution, obey an inverse power law with respect to their power spectral density [51–53]. Specifically, the power spectrum of a particular image drops off a $1/f^{\gamma}$, with γ on the order of three. This indicates that the images are predominately low-frequency fields with long-range, although not well defined, spatial correlation. Power laws are inherently termed self-similar and often the term fractal is used. This implies that there are no preferred scales as with the human voice for example. There are debates as to whether an anatomical structure like the breast could be truly fractal. But, it is reasonable to say that the image statistics will quite often vary widely across the image from region-to-region due to this spectral characteristic. As an aside, wavelet expansion images may be considered as multiresolution derivatives (derivatives with respect to scale) in a loose sense. Effectively, the differencing produces images (the expansion image) that are not as irregular as the raw images. The reader interested in this line of reasoning could consult Heine *et al.* [52, 54] and the references therein.

13.3.4 Adaptive Thresholding

Following the filtering approaches described in sections 13.3.3 and 13.3.4, calcifications were segmented by either an adaptive thresholding approach or a Canny edge detector. The former method yielded better classification results so far with either the symmlet wavelet or the donut filter and this will be discussed here in more detail. Figures 13.13 and 13.14 show the results of the thresholding process applied on the filter outputs of Figs. 13.10 and 13.11.

To reduce FP signals in either output, a criterion was set on the minimum size of the segmented objects based on empirical observations of calcifications

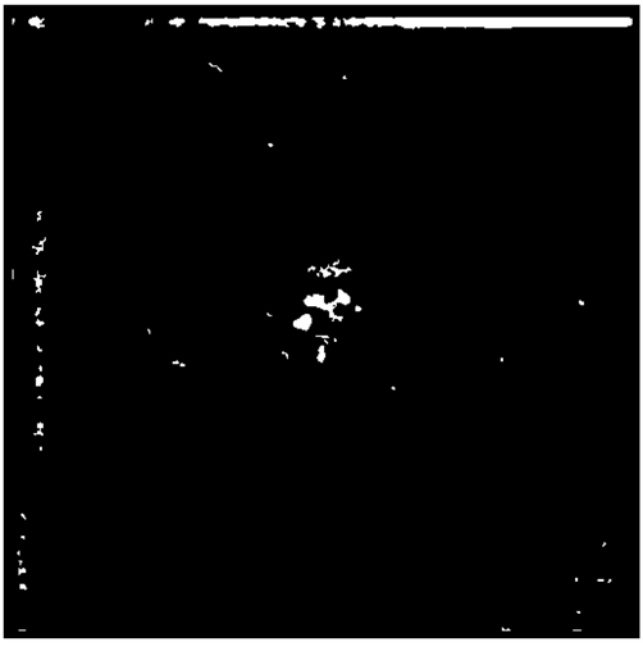


Figure 13.13: Adaptive thresholding of the symmlet wavelet filter's output of Fig. 13.10. The true calcifications are isolated in addition to false signals generated by calcified arteries or tissue intersections that "look like" calcification structures and have similar spectral properties. The edge effects shown in Fig. 13.10 remain as white borders in this stage that can be removed at the expense of losing details in calcification morphology, size, and number particularly for very small calcifications.

and visibility limits reported for calcifications in mammography literature [55]. Specifically, segmented spots smaller than 4 pixels (0.0144 mm^2) in area, of any configuration, were eliminated from the final segmentation step prior to shape analysis and classification.

13.4 Shape Analysis and Classification Feature Definition

According to the flowchart of Fig. 13.3, the steps following the detection and segmentation of the calcifications involve shape analysis of the segmented

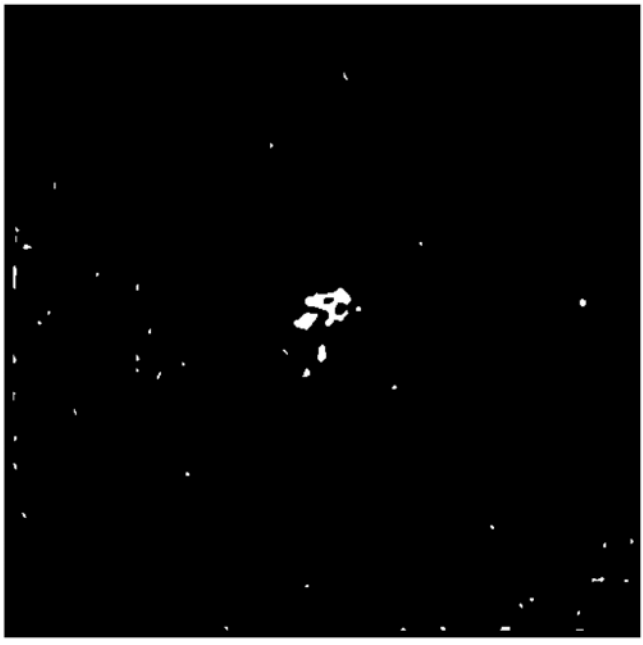


Figure 13.14: Adaptive thresholding of the donut filter output of Fig. 13.11. As in Fig. 13.13, both true and false calcifications were isolated and outlined. No edge effects were generated in this case. Furthermore, more calcifications were preserved in the segmentation stage at the expense of a slightly higher number of false signals.

objects and selection of the feature set to be used as input to the classifier. For this stage, we took advantage of prior art in the field of classification and our experience in mammographic features [56]. Our starting point was the implementation of the four shape features developed by Shen *et al.* [57] for individual calcifications and their modification to apply to calcification clusters. We expanded this initial set with two more shape descriptors of individual calcifications [20]. To represent the clusters, we added the standard deviations of the six shape descriptors and a distribution feature. To represent the patient and link the demographic data to the images, we added a demographic feature [58]. The final results was a set of fourteen features for cluster classification in mammography. Table 13.2 lists the selected feature set and the physical interpretation of each feature [59]. Specific definitions and details may be found in the listed references.

Table 13.2: Feature set selected from the shape analysis of the segmented individual calcifications and clusters and demographic data[a]

Feature No.	Feature	Nature of feature
1	Age of the patient	*Demographic* feature; describes the patient
Individual calcification characteristics		
2	Mean—Area of calcification	Describes the *morphology* (shape)
3	Mean—Compactness	
4	Mean—Moments	
5	Mean—Fourier Descriptor (FD)	Describes the *margins*
6	Mean—Eccentricity	
7	Mean—Spread (S)	
8	Number of calcifications in cluster (median of range)	*Regional* descriptor; describes distribution
Cluster characteristics		
9	SD—Area	Describes the *morphology*
10	SD—Compactness	
11	SD—Moments	
12	SD—Fourier Descriptor	Describes the *margins*
13	SD—Eccentricity	
14	SD—Spread	

[a]Features are limited to morphological and distributional characteristics (with the exception of "age") in order to reproduce the visual analysis system and indirectly use the classification as a measure of segmentation.

13.5 Classification Algorithm

Classification was done with a three-layer, feed-forward artificial neural network (ANN) consisting of an input layer, one hidden layer, and an output layer. The NevProp4 backpropagation software was used in this study. NevProp4 is a general backpropagation algorithm developed by Philip H. Goodman at the University of Nevada, Reno [60]. Figure 13.15 shows a diagram of the network structure.

The feature vector of the input layer consisted of 14 elements (features) defined in the previous stage (Table 13.2) and a bias element [60]. The hidden layer consisted of 12 nodes and the output layer had one node. For each cluster, the network was given the set of shape features at its input layer, merged these inputs internally using the hidden and output layers, and assigned a value in the range of 0–1, where 0 was the target output for the benign cases and 1 was the

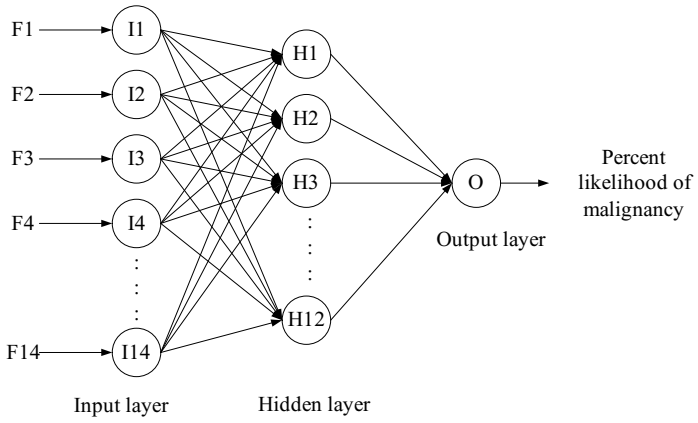


Figure 13.15: Diagram of the NevProp4 artificial neural network (ANN) used for cluster classification. This is a standard three-layer, feed-forward ANN where F1–F14 are the input features, I1–I14 are the input units, H1–H12 are the hidden units, and O is the output layer [20, 59, 60].

target output for the cancer cases. This value could be interpreted as a percent likelihood for a cluster to be malignant.

The generalization error of the ANN classifier was estimated by the "leave-one-out" resampling method [61, 62]. Leave-one-out is a method generally recommended for the validation of pattern recognition algorithms using small datasets. The use of this approach usually leads to a more realistic index of performance and eliminates database problems such as small size and not fully representative contents and problems associated with the mixing of training and testing datasets [61, 63]. In the leave-one-out validation process, the network was trained on all but one of the cases in the set for a fixed number of iterations and then tested on the one excluded case. The excluded case was then replaced, the network weights were reinitialized, and the training was repeated by excluding a different case until every case had been excluded once. For N cases, each exclusion of one case resulted in N–1 training cases, 1 testing case and a unique set of network weights. As the process was repeated over all N, there were $N(N-1)$ training outputs and N testing outputs from which the training and testing mean square error (MSE) was, respectively, determined.

In addition to the leave-one-out method, other resampling approaches have been proposed for CADiagnosis algorithm training that could yield unbiased

results and provide meaningful and realistic estimates of performance. A preference toward the bootstrap technique is found in the literature although this is strongly dependent on the application and availability of resources [64]. There is considerable work reported in the field and we will not elaborate more in this chapter. The reader, however, should be aware of the bias issues associated with large feature sets and small sample sizes and the possible methods of training and testing an algorithm. An approach should generally be selected that yields no overestimates of performance and avoids training bias.

The clinical value of CADiagnosis methods is usually assessed in two stages: First, computer performance is evaluated based on truth files defined by the experts and biopsy information using computer generated receiver operating characteristic (ROC) curves [65, 66]. Computer ROC is implemented in the evaluation of classification algorithms where sensitivity and specificity indices are generated by adjusting the algorithms' parameters. Classification algorithms differentiate usually between benign vs. cancer lesions, disease vs. not disease, disease type 1 vs. disease type 2, etc. The pairs of sensitivity and specificity generated by these algorithms can be plotted as a true positive fraction (TPF) vs. false positive fraction (FPF) to form an ROC curve [65]. Publicly available software tools, e.g., the ROCKIT from Metz at the University of Chicago [67], may be used to fit the data and estimate performance parameters such as the area under the curve, A_Z, its standard error (SE), confidence intervals, and statistical significance.

Following the laboratory evaluation, a true ROC experiment is usually performed that involves relatively large number of cases and human observers [68]. The cost and time requirements of an observer ROC study are significant impediments in its implementation and such analysis is usually reserved for fully optimized techniques, namely for techniques that have been through rigorous computer ROC evaluation. Computer ROC evaluation poses specific requirements on database size and contents and the criteria used for the estimation of TPF and FPF values at the detection or the classification level. We will not labor on these issues in this chapter; guidelines may be found in several publications in the field of CAD and elsewhere [65, 69, 70]. We will only mention that a sufficiently large set should be selected for CADiagnosis validation to meet the requirements of the classification scheme, while the contents of the dataset should be such as to address the specific clinical goals of the methodology. In addition, performance criteria should follow clinical guidelines and

be consistently and uniformly applied throughout the validation process. In the CADiagnosis algorithm applications presented below, equal numbers of benign and malignant cases with calcification clusters were used and almost all cluster shapes described in the BIRADS Lexicon [71] were represented in the sets. Performance parameters such as number of TP and FP clusters at the segmentation output or TPF and FPF at the classification output were estimated based on well-defined criteria that were consistently applied to all experiments.

13.6 CADiagnosis Algorithm Applications

Several applications of the algorithm described here have been reported in the literature [20, 28, 29]. In this chapter, we summarize the most important ones, report on new experiments that are linked to segmentation issues and reveal some of the open questions remaining in this area.

13.6.1 Mammographic Cluster Classification—Single View Application

A set of 260 single-view mammograms with calcification clusters was first used for the validation of the CADiagnosis algorithm described previously. The set included 138 calcification clusters associated with benign disease that are commonly referred to as benign calcifications or clusters and 122 calcification clusters associated with cancer that are commonly referred to as malignant calcifications or clusters. All mammograms were selected from the patient files of the H. Lee Moffitt Cancer Center & Research Institute at the University of South Florida. Original mammograms were acquired on two different mammography systems, both accredited by the American College of Radiology (ACR) and having similar performances. A DuPont Microvision film combined with a Kodak Min-R (one-sided) screen was used for all mammograms. Films were digitized with a DBA (DBA Inc., Melbourne, FL) ImagClear R3000 CCD-based film digitizer with a pixel size of 30 μm, a pixel depth of 16 bits, and a nonlinear response to optical density [72]. Full images were resized to 60 μm by mathematical interpolation keeping the pixel depth the same. For this application, 512×512 pixel ROIs were processed. ROIs were selected from the full 60 μm images to contain the calcification cluster of interest.

Mammographic views were either cranio-caudal (CC) or medio-lateral oblique (MLO) views of the right of left breast. Two hundred two (202) views from this set were images of the same cluster that is they were CC and MLO views of the same breast and the same patient. For this application, however, they were considered as independent samples. This is common practice in the field not only because of the rarity of the data but also because most CADetection schemes today are applied to single views only and do not usually consider the full mammogram or the "other" breast view in the process. A bias is certainly expected when views from the same patient and of the same cluster are treated as independent samples and this bias could affect performance. This is investigated in the following application example described in section 13.6.2.

The CADiagnosis scheme of Fig. 13.3 was applied to the set of 260 ROIs first using the symmlet wavelet filter and then using the donut filter in the detection/segmentation stage. The computer ROC curves obtained from the two classification experiments are shown in Fig. 13.16. The corresponding A_Z performance index values and standard errors (SE) are included on the figure. The difference between the two curves was statistically significant indicat-

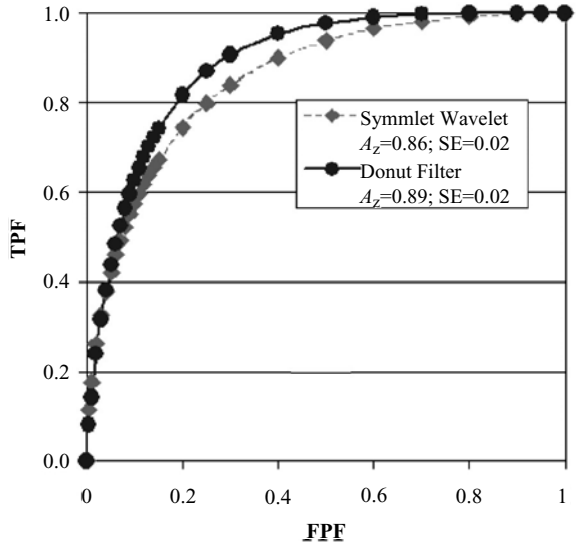


Figure 13.16: Computer ROC plots of the TPF and FPF pairs obtained from the classification of 260 clusters. The dashed curve corresponds to the results obtained with the symmlet wavelet filter and the solid curve corresponds to the results obtained with the donut filter. The estimated area indices A_Z and corresponding SE values are included in the insert.

ing that cluster classification using the donut filter in the detection and segmentation stage was significantly better than classification using the symmlet wavelet.

13.6.2 Mammographic Cluster Classification—Two-View Application

In this application of the algorithm, the two views of the same cluster were combined for the selection of the classification features [59]. A total of 101 paired clusters were available for this test. The 14 features of Table 13.2 were first determined on the 101 CC and 101 MLO views of the cluster and then averaged. The set of average feature values were then used as input to the classification stage of the algorithm (Fig. 13.3). The computer generated ROC curves of the classification performance of the algorithm obtained with the symmlet wavelet and the donut filter are shown in Fig. 13.17. Similar to the previous experiment, the classifier with the donut filter outperformed the classifier with the

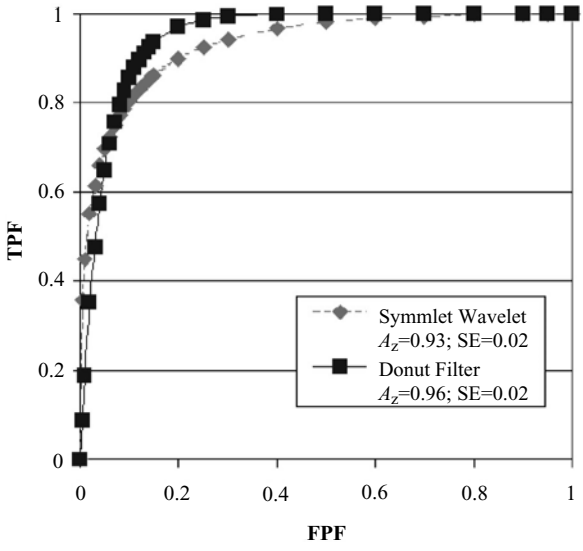


Figure 13.17: Computer ROC plots of the TPF and FPF pairs obtained from the classification of 101 clusters using two-view information for feature estimation. The dashed curve corresponds to the results obtained with the symmlet wavelet filter and the solid curve corresponds to the results obtained with the donut filter. The estimated area indices A_Z and corresponding SE values are included in the insert.

symmlet wavelet. And both outperformed their respective performances on the single-view application. The results suggest that the combination of views for feature estimation seems beneficial to the classification process.

Two views could lead to the definition of more robust features improving classification performance independent of the segmentation method used in the process. But, is averaging the best approach to feature selection from the two mammographic views? Our results seem to indicate that averaging is promising. However, they are somewhat counterintuitive since averaging carries the risk of introducing a fuzziness to an otherwise good descriptor, i.e., a feature that was a good descriptor in one view but poor in the other may lose its robustness once averaged. So, should we average or possibly combine features from the two views for the generation of a larger feature set? The answer to this question is not clear and more work is needed to determine the best feature combination from two mammographic views.

13.6.3 Impact of False Positive Segmentations on Classification

In this third experiment, we examined the impact of the FP segmentations on the performance of the classifier. As seen in Figs. 13.13 and 13.14, several signals remain in the segmentation output that are not true calcifications, individual, or clusters. These signals enter the classification stage of the algorithm and are likely to affect performance. To determine the degree of this effect, we first estimated the number of TP and FP segmented clusters. This was done be comparing the segmentation output to manual outlines of the clusters and their major calcifications generated by expert mammographers. The guidelines and conventions described elsewhere [70] were followed for these estimates. Specifically, a segmented group of calcifications was considered a TP when it contained at least three segmented true calcifications [71]. A FP cluster was one that consisted of at least three segmented objects outside the area of the true cluster within a distance of $\leq$1 cm from each other. Following the above guidelines, we determined that for a 100% TP rate, an average of 2.8 FP clusters were segmented per image with the symmlet wavelet and an average of 2.0 FP clusters were segmented with the donut filter. A reduction in either FP rate was always followed by a reduction in the TP rate to levels that were not acceptable

by the classification stage that, in our case, is heavily based on morphology and distribution characteristics of the individual calcifications and their clusters.

To study the impact of the FP signals on performance but without losing TP information, we did the following experiments:

(a) The 512×512 pixel ROIs of the 260 clusters were automatically reduced to 200×200 pixels. As a result, several of the edge effects and associated false signals were eliminated concentrating the analysis on the center of the region where the signal of interest (cluster) should normally be present. Both algorithm versions were applied to the reduced-size ROIs. Results suggested that the classification of both the benign and malignant cases might be improved by up to 15% for the algorithm using the symmlet wavelet filter and up to 10% for the algorithm using the donut filter. The smaller improvement in the latter case was expected because the donut filter did not show major edge effects as the symmlet wavelet did in the original ROIs (see Figs. 13.13 and 13.14). This seemed to be an easy and fast remedy to the problem of FP signals with one downside. Namely, if the clusters were off-center in the initial ROI either due to their physical location in the breast (e.g., close to the chest wall or the skin area) or due to the initial ROI selection, then important information was lost and classification could not be done successfully.

(b) In a second experiment, all FP clusters and all single, isolated false signals that were outside the boundaries of the true cluster were manually eliminated from the 512×512 pixel ROIs. This manual elimination was done for a subset of 30 cases that contained small calcification clusters (3–10 calcifications per cluster). The original and FP-free ROIs were then used for feature estimation and classification. The elimination of the FP signals improved the classification of both benign and malignant cases. Significant classification improvement was observed for both benign and malignant calcification clusters and both algorithm versions. Classification errors were reduced up to 30% for the benign cases and up to 50% for the malignant cases. Further analysis of these results revealed that the presence of very small false objects in the segmentation output degraded classification performance more than large false objects such as those originating from the edge artifacts.

13.7 Conclusions

CADiagnosis is an area that merges the fields of signal processing and pattern recognition for the creation of tools that can have a significant impact in health care delivery and patient management. CADiagnosis algorithms usually involve several modules that need to be separately optimized and validated for an overall optimum performance. In this chapter, we presented a CADiagnosis methodology for the differentiation between benign and malignant breast calcification clusters in mammograms. We specifically looked into one of the aspects of the algorithm, namely the impact of segmentation in the overall classification process, and the role of multiresolution analysis in the segmentation process.

Our classification approach was based primarily on morphological and distributional features of mammographic calcifications and, hence, the role of segmentation was particularly important in the overall implementation and performance. Knowing the limitations of image segmentation techniques that were further exaggerated by our additional challenge to preserve morphology and distribution, we developed two multiresolution filters that were able to yield successful and clinically promising results. Although far from perfect segmentations, the symmlet wavelet and the donut filter adequately preserved the characteristics of the calcifications as required by the overall algorithm's design. A new filter, labeled as the "donut filter," was introduced for mammogram processing that seems to offer a robust solution to the problems associated with the detection and segmentation of mammographic images. The new filter was not utilized to its full potential and several implementation pathways remain to be explored. Its initial testing, however, yielded promising results and its usefulness could go beyond mammography to other medical imaging applications.

An important question at the end of the experiments presented here is whether similar classification performance can be achieved, either with the symmlet wavelet or the donut filter, for images generated from various sources. For example, for images generated by different film digitizers (laser-based vs. charge-couple-device-based systems), or by different imaging systems (screen/film vs. direct digital systems), or with different resolution characteristics (pixel size and bit depth). Preliminary work with different data types suggests that similar classification results may be obtained if a standardization process is applied to the images prior to processing. As long as pixel size and depth are within acceptable ranges for CADiagnosis applications in mammography,

a standardization algorithm can easily convert the characteristics of any set of data to those for which the CADiagnosis system was initially trained and optimized keeping performance consistent [20, 73].

An interesting spin-off application of our initial development originated from the FP impact observation on classification performance. We found that classification results could be used as an indirect measure of segmentation quality particularly when the classification scheme is based solely on morphological and distributional characteristics like the one described here. Segmentation evaluation is one of the most challenging issues in medical image processing. It usually requires objective and accurate "ground truth" or "gold standard" information that is often unattainable in medical imaging where the human observer is commonly the only source of "ground truth" information. Using the classifier's output for indirect segmentation validation may offer an advantage over more traditional techniques that use absolute measures of shape and size and require exact ground truth information. After all, it is the clinical outcome that is important in these applications.

Finally, the described CADiagnosis scheme seems to be amendable to a variety of applications beyond breast cancer screening and early diagnosis. The input feature set and classification output could be modified and expanded to address problems associated with the diagnostic patient and specific breast disease types involving calcifications, e.g., ductal carcinoma in-situ, for the development of computer tools that go beyond detection and diagnosis into the domains of prognosis, patient management, and follow-up.

13.8 Acknowledgments

The authors acknowledge the valuable assistance of Angela Salem in the generation of the image databases used for algorithm development and testing, and of Joseph Murphy in the processing of the data.

Questions

1. *Role of CADetection and CADiagnosis techniques in breast cancer and mammography. What is the relationship of the two systems?*

2. *What are the basic elements and characteristics of a CADiagnosis scheme?*
3. *What is the "visual analysis" system of mammographic calcifications and how can it be translated to a computer methodology that helps mammogram interpretation?*
4. *What property exists between wavelet expansion images that does not exist in an arbitrary filter bank output?*
5. *What can you say about a time series signal that is nothing but a spike a* $t = t_0$ *with respect to its Fourier properties?*
6. *An image with long-range positive correlation will have large Fourier components in what part of the frequency domain?*
7. *Explain what a band pass filter is and what it may be used for.*
8. *What is white noise and give an example from common observation?*
9. *Given a low-frequency signal of interest buried in white noise, what kind of filter would work for lessening the influence of the noise?*
10. *What are the criteria for database design as needed for the evaluation of CADiagnosis algorithms?*
11. *What are the validation steps of a CADiagnosis scheme?*
12. *What is the different between computer ROC and observer ROC? How can we correlate the ROC indices of performance to clinically used indices of sensitivity and specificity?*
13. *Could segmentation be validated through classification and how?*

Bibliography

[1] Huo, Z., Giger, M. L., Vyborny, C. J., and Metz, C. E., Breast cancer: Effectiveness of computer-aided diagnosis-observer study with independent database of mammograms, Radiology, Vol. 224, pp. 560–568, 2002.

[2] Feig, S. A., Clinical evaluation of computer-aided detection in breast cancer screening, Sem. Breast Dis., Vol. 5, No. 4, pp. 223–230, 2002.

[3] de Koning, H. J., Mammographic screening: Evidence from randomized controlled trials, Ann Oncol., Vol. 14, No. 8, pp. 1185–1189, 2003.

[4] Feig, S. A., Decreased breast cancer mortality through mammographic screening: Results of clinical trials, Radiology, Vol. 167, pp. 659–665, 1988.

[5] Clark, R. A., Breast cancer screening: Is it worthwhile? Cancer Control, Vol. 3, pp. 189–194, 1995.

[6] Bird, R. E., Wallace, T., W., and Yankaskas, B. C., Analysis of cancers missed at screening mammography, Radiology, Vol. 184, No. 3, pp. 613–617, 1992.

[7] Millis, R. R., Davis, R., and Stacey, A. J., The detection and significance of calcifications in the breast: A radiological and pathological study, Br. J. Radiol., Vol. 49, pp. 12–26, 1976.

[8] Reintgen, D., Berman, C., Cox, C., Baekey, P., Nicosia, S., Greenberg, H., Bush, C., Lyman, G. H., and Clark, R. A., The anatomy of missed breast cancer, Surg. Oncol., Vol. 2, pp. 65–75, 1993.

[9] Kopans, D. B., The positive predictive value of mammography, AJR, Vol. 158, No. 3, pp. 521–526, 1992.

[10] Lewin, J. M., Hendrick, R. E., D'Orsi, C. J., Isaacs, P. K., Moss, L. J., Karellas, A., Sisney, G. A., Kuni, C. C., and Cutter, G. R., Comparison of full-field digital mammography with screen-film mammography for cancer detection: Results of 4,945 paired examinations, Radiology, Vol. 218, pp. 873–880, 2001.

[11] Giger, M. L., Computer-aided diagnosis in radiology, Acad. Radiol., Vol. 9, No. 1, pp. 1–3, 2002.

[12] Nishikawa, R., Assessment of the performance of computer-aided detection and computer-aided diagnosis systems, Sem. Breast Dis., Vol. 5, No. 4, pp. 217–222, 2002.

[13] Floyd, C. E., Lo, J. Y., Yun, A. J., Sullivan, D. C., and Kornguth, P. J., Prediction of breast cancer malignancy using an artificial neural network, Cancer, Vol. 74, No. 11, pp. 2944–2948, 1994.

[14] Jiang, Y., Nishikawa, R. M., Schmidt, R. A., Metz, C. E., Giger, M. L., and Doi, K., Improving breast cancer diagnosis with computer-aided diagnosis, Acad. Radiol., Vol. 6, pp. 22–33, 1999.

[15] Giger, M. L., Huo, Z., Kupinski, M. A., and Vyborny, C. J., Computer-aided diagnosis in mammography, In: Handbook of Medical Imaging, Volume 2, Medical Image Processing and Analysis, Sonka, M. and Fitzpatrick, M. J., eds., SPIE Press, Bellingham, WA, pp. 915–1004, 2000.

[16] Li, L., Zheng, Y., Zhang, L., and Clark, R. A., False-positive reduction in CAD mass detection using a competitive strategy, Med. Phys., Vol. 28, No. 2, pp. 250–258, 2001.

[17] Wu, Y., Giger, M. L., Doi, K., Vyborny, C. J., Schmidt, R. A., and Metz, C. E., Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer, Radiology, Vol. 187, pp. 81–87, 1993.

[18] Chan, H. P., Sahiner, B., Kam, K. L., Petrick, N., Helvie, M. A., Goodsitt, M. M., and Adler, D. D., Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces, Med. Phys., Vol. 25, No. 10, pp. 2007–2019, 1998.

[19] Jiang, Y., Nishikawa, R. M., Wolverton, D. E., Metz, C. E., Giger, M. L., Schmidt, R. A., Vyborny, C. J., and Doi, K., Malignant and benign clustered microcalcifications: Automated feature analysis and classification, Radiology, Vol. 198, pp. 671–678, 1996.

[20] Kallergi, M., Computer aided diagnosis of mammographic microcalcification clusters, Med. Phys., Vol. 31, pp. 314–326, 2004.

[21] Lanyi, M., Morphologic analysis of microcalcifications: A valuable differential diagnostic system for early detection of breast carcinomas and reduction of superfluous exploratory excisions, In: Early Breast Cancer, Zander, J. and Baltzer, J., eds., Springer-Verlag, Berlin, pp. 113–135, 1985.

[22] Lanyi, M., Diagnosis and Differential Diagnosis of Breast Calcifications, Springer-Verlag, Berlin, 1986.

[23] Hall, F. M., Storella, J. M., Silverstone, D. Z., and Wyshak, G., Nonpalpable breast lesions: Recommendations for biopsy based on suspicion of carcinoma at mammography, Radiology, Vol. 167, pp. 353–358, 1988.

[24] Olson, S. L., Fam, B. W., Winter, P. F., Scholz, F. J., Lee, A. K., and Gordon, S. E., Breast calcifications: Analysis of imaging properties, Radiology, Vol. 169, pp. 329–332, 1988.

[25] Muir, B. B., Lamb, J., Anderson, T. J., and Kirkpatrick, A. E., Microcalcification and its relationship to cancer of the breast: Experience in a screening clinic, Clin. Radiol., Vol. 34, pp. 193–200, 1983.

[26] D'Orsi, C. J. and Kopans, D. B., Mammographic feature analysis, Sem. Roentgenol., Vol. XXVIII, No. 3, pp. 204–230, 1993.

[27] Liberman, L., Abramson, A. F., Squires, F. B., Glassman, J. R., Morris, E. A., and Dershaw, D. D., The breast imaging reporting and data system: Positive predictive value of mammographic features and final assessment categories, AJR, Vol. 171, pp. 35–40, 1998.

[28] Kallergi, M., Gavrielides, M. A., He, L., Berman, C. G., Kim, J. J., and Clark, R. A., A simulation model of mammographic calcifications based on the ACR BIRADS, Acad. Radiol., Vol. 5, pp. 670–679, 1998.

[29] Gavrielides, M. A., Kallergi, M., and Clarke, L. P., Automatic shape analysis and classification of mammographic calcifications, In: SPIE, Vol. 3034, pp. 869–876, 1997.

[30] Tolstov, G. P., Fourier Series, Dover Publications, New York, 1962.

[31] Bracewell, R. L., The Fourier Transform and Its Applications, 2nd edn. revised, McGraw-Hill, New York, 1988.

[32] Brigham, E. O., The Fast Fourier Transform and Its Applications, Prentice Hall, Englewood Cliffs, NJ, 1988.

[33] Bracewell, R. L., Two-Dimensional Imaging, Prentice Hall, Englewood Cliffs, NJ, 1995.

[34] Beckmann, P., Probability in Communication Engineering, Harcort, Brace & World, New York, 1967.

[35] Thomas, J. B., An Introduction to Statistical Communication Theory, Wiley, New York, 1969.

[36] Helstrom, C. W., Probability and Stochastic Processes For Engineers, 2nd edn., Macmillan, New York, 1991.

[37] Papoulis, A., Probability, Random Variables, and Stochastic Processes, 3rd edn., McGraw-Hill, Boston, MA, 1991.

[38] Giffin, W. C., Transform Techniques for Probability Modeling, Academic Press, New York, 1975.

[39] Strang, G. and Nguyen, T., Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996.

[40] Akansu, A. N. and Haddad, R. A., Multiresolution Signal Decomposition Transforms, Subbands, and Wavelets, Academic Press, Boston, MA, 1992.

[41] Vetterli, M. and Kovacevic, J., Wavelets and Subband Coding, Prentice Hall, Englewood Cliffs, NJ, 1995.

[42] Daubechies, I., Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[43] Peitgen, H. O., Jurgens, H., and Saupe, D., Chaos and Fractals: New Frontiers of Science, Springer-Verlag, New York, 1992.

[44] Wornell, G. W., Signal Processing with Fractals: A Wavelet Based Approach, Prentice Hall, Upper Saddle River, NJ, 1996.

[45] Turner, M. J., Blackledge, J. M., and Andrews, P. R., Fractal Geometry in Digital Imaging, Academic Press, San Diego, CA, 1998.

[46] Heine, J. J., Deans, S. R., Cullers, D. K., Stauduhar, R., and Clarke, L. P., Multiresolution statistical analysis of high-resolution digital mammograms, IEEE Trans. Med. Imag., Vol. 16, No. 5, pp. 503–604, 1997.

[47] Heine, J. J., Deans, S. R., and Cullers, D. K., Stauduhar, R., and Clarke, L. P., Multiresolution probability analysis of gray scaled images, J. Opt. Soc. Am. A, Vol. 15, pp. 1048–1058, 1998.

[48] Heine, J. J., Deans, S. R., and Clarke, L. P., Multiresolution probability analysis of random fields, J. Opt. Soc. Am. A, Vol. 16, pp. 6–16, 1999.

[49] Mendenhall, W. and Scheaffer, R. L., Mathematical Statistics with Applications, Duxbury Press, North Scituate, MA, 1973.

[50] D'Orsi, C. J. and Kopans, D. B., Mammographic feature analysis, Sem. Roentgenol., Vol. XXVIII, No. 3, pp. 204–230, 1993.

[51] Heine, J. J., Multiresolution statistical analysis of direct x-ray detection digital mammograms, Final report, Department of Defense, CDMRD, 2002.

[52] Heine, J. J., Deans, S. R., Velthuizen, R. P., and Clarke, L. P., On the statistical nature Of mammograms, Med. Phys., Vol. 26, pp. 2254–2265, 1999.

[53] Burgess, A. E., Jacobson, F. L., and Judy, P. F., Human observer detection experiments with mammograms and power-law noise, Med. Phys., Vol. 28, No. 4, pp. 419–437, 2001.

[54] Heine, J. J., Deans, S. R., Gangadharan, D., and Clarke, L. P., Multiresolution analysis of two dimensional 1/f processes: Approximations, Opt. Eng., Vol. 38, pp. 1505–1516, 1999.

[55] Freedman, M., Pe, E., Zuurbier, R., Katial, R., Jafroudi, H., Nelson, M., Lo, S. C. B., and Mun, S. K., Image processing in digital mammography, SPIE, Vol. 2164, pp. 537–554, 1994.

[56] Woods, K., Automated Image Analysis Techniques for Digital Mammography, Ph.D. Dissertation, Department of Computer Science and Engineering, College of Engineering, University of South Florida, 1994.

[57] Shen, L., Rangayyan, R. M., and Desautels, J. E. L., Application of shape analysis to mammographic calcifications, IEEE Trans. Med. Imag., Vol. 13, No. 2, pp. 263–274, 1994.

[58] Jemal, A., Thomas, A., Murray, T., and Thun, M., Cancer statistics 2002, CA Cancer J. Clin., Vol. 52, pp. 23–47, 2002.

[59] Tembey, M., Computer Aided Diagnosis for Mammographic Microcalcification Clusters, MS Thesis, Computer Science Department, College of Engineering, University of South Florida, Tampa, FL, 2003.

[60] Burke, H. B., Goodman, P. H., Rosen, D. B., Henson, D. E., Weinstein, J. N., Harrell, F. E., Marks, J. R., Winchester, D. P., and Bostwick, D. G., Artificial neural networks improve the accuracy of cancer survival prediction, Cancer, Vol. 79, pp. 857–862, 1997.

[61] Efron, B., The Jacknife, the Bootstrap, and Other Resampling Plans, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982.

[62] Tourassi, G. D. and Floyd, C. E., The effect of data sampling on the performance evaluation of artificial neural networks in medical diagnosis, Med. Decis. Making, Vol. 17, pp. 186–192, 1997.

[63] Harrell, F. E., Lee, K. L., and Mark, D. B., Tutorial in biostatistics, multivariate prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Stat. Med., Vol. 15, pp. 361–387, 1996.

[64] Chen, D. R., Kuo, W. J., Chang, R. F., Moon, W. K., and Lee, C. C., Use of the bootstrap technique with small training sets for computer-aided diagnosis in breast ultrasound, Ultrasound Med. Biol., Vol. 28, No. 7, pp. 897–902, 2002.

[65] Nishikawa, R., Assessment of the performance of computer-aided detection and computer-aided diagnosis systems, Sem. Breast Dis., Vol. 5, No. 4, pp. 217–222, 2002.

[66] Bowyer, K. W., Validation of medical image analysis techniques, In: Handbook of Medical Imaging, Volume 2, Medical Image Processing and Analysis, Sonka, M. and Fitzpatrick, J. M., eds., SPIE Press, Bellingham, WA, pp. 567–607, 2000.

[67] University of Chicago. Kurt Rossmann Laboratories for Radiologic Image Research. http://home.uchicago.edu/njunji/KRL_HP/roc_soft.htm. Accessed September 2, 2004.

[68] Dorfman, D. D., Berbaum, K. S., and Lenth R. V., Multireader, multicase receiver operating characteristic methodology: A bootstrap analysis, Acad. Radiol., Vol 2, pp. 626–633, 1995.

[69] Nishikawa, R. M., Giger, M. L., Doi, K., Metz, C. E., Yin, F. F., Vyborny, C. J., and Schmidt R. A., Effect of case selection on the performance of computer-aided detection schemes, Med. Phys., Vol. 21, No. 2, pp. 265–269, 1994.

[70] Kallergi, M., Carney, G., and Gaviria, J., Evaluating the performance of detection algorithms in digital mammography, Med. Phys., Vol. 26, No. 2, pp. 267–275, 1999.

[71] D'Orsi, C. J. and Kopans, D. B., Mammographic feature analysis, Sem. Roentgenol., Vol. XXVIII, No. 3, pp. 204–230, 1993.

[72] Kallergi, M., Gavrielides, M. A., Gross, W. W., and Clarke, L. P., Evaluation of a CCD-based film digitizer for digital mammography, SPIE, Vol. 3032, pp. 282–291, 1997.

[73] Velthuizen, R. P. and Clarke, L. P., Digitized mammogram standardization for display and CAD, SPIE, Vol. 3335, pp. 179–187, 1998.

Chapter 14

Computer-Supported Segmentation of Radiological Data

Philippe Cattin,[1] *Matthias Harders*,[1] *Johannes Hug*,[1]
Raimundo Sierra,[1] *and Gabor Szekely*[1]

14.1 Introduction

Segmentation is in many cases the bottleneck when trying to use radiological image data in many clinically important applications as radiological diagnosis, monitoring, radiotherapy, and surgical planning. The availability of efficient segmentation methods is a critical issue especially in the case of large 3-D medical datasets as obtained today by the routine use of 3-D imaging methods like magnetic resonance imaging (MRI), computer tomography (CT), and ultrasound (US).

Although manual image segmentation is often regarded as a gold standard, its usage is not acceptable in some clinical situations. In some applications such as computer-assisted neurosurgery or radiotherapy planning, e.g., a large number of organs have to be identified in the radiological datasets. While a careful and time-consuming analysis may be acceptable for outlining complex pathological objects, no real justification for such a procedure can be found for the delineation of normal, healthy organs at risk. Delineation of organ boundaries is also necessary in various types of clinical studies, where the correlation between morphological changes and therapeutical actions or clinical diagnosis has to be analyzed. In order to get statistically significant results, a large number of datasets has to be segmented. For such applications manual segmentation

[1] Computer Vision Laboratory, ETH-Zurich, Switzerland

becomes questionable not only because of the amount of work, but also with regard to the poor reproducibility of the results.

Because of the above reasons, computer-assisted segmentation is a very important problem to be solved in medical image analysis. During the past decades a huge body of literature has emerged, addressing all facets of the related scientific and algorithmic problems. A reasonably comprehensive review of all relevant efforts is clearly beyond the scope of this chapter. Instead, we just tried to analyze the underlying problems and principles and concisely summarize the most important research results, which have been achieved by several generations of PhD students at the Computer Vision Laboratory of the Swiss Federal Institute of Technology during the past 20 years.

14.2 Intensity-Based Automatic Segmentation

Early approaches for automatic segmentation fundamentally use the assumption that radiological images are basically "self-contained," i.e., they contain most of the information which is necessary for the identification of anatomical objects. In some limited applications such techniques can be very successful, as the automatic segmentation of dual-echo MR images [1]. This example will be used here as an illustration as it addresses most aspects of intensity-based medical image segmentation. The method uses two spatially perfectly matched echos of a spin-echo MR acquisition as illustrated by Figs. 14.1(a) and 14.1(c).

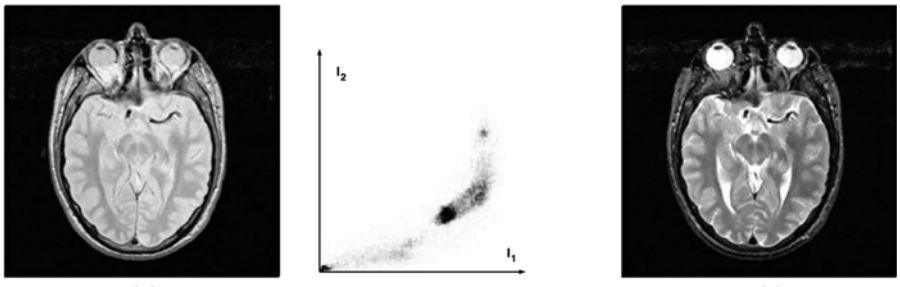
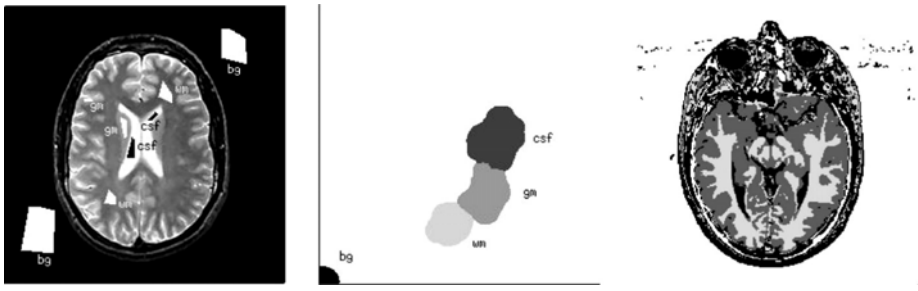


Figure 14.1: Spin-echo MR image pair (an early echo is shown on the left, a late echo on the right). In the middle the two-dimensional intensity distribution (i.e., the frequency of the occurrence of intensities I_1 and I_2 in the left and right images) is given.



(a) **(b)** **(c)**

Figure 14.2: Segmentation of the dual-echo MR image using training. The left image shows user-defined training regions for the different tissue classes. The corresponding tessellation of the feature space (spanned by the joint intensity distribution) is shown in the middle, resulting in the segmentation on the right.

The applied procedure can be regarded as a generalized thresholding, aiming at the identification of areas in a feature space, i.e. in the two-dimensional intensity distribution (Fig. 14.1(b)), which uniquely characterize the different tissue classes (as gray or white matter of the brain). These areas are usually determined during a training phase, where the user identifies examples for each tissue class (e.g. in the form of regions of interest as illustrated in Fig. 14.2(a)). Standard pattern recognition procedures (e.g., as k-nearest neighbor classification) [2] can be used to derive a corresponding tessellation of the feature space (Fig. 14.2(b)) leading to the classification of the entire image (Fig. 14.2(c)).

The success of the segmentation basically depends on the assumption that tissue classes can perfectly be separated in the feature space provided by the measurements. Beside physiologically induced overlaps between features of different tissue classes, limitations of the acquisition process can seriously compromise the efficiency of the method.

The most important sources of error are the presence of noise, the spatial inhomogeneity of the signal intensity generated by the tissue, and the limited spatial resolution of the images leading to partial volume effects.

The presence of voxels containing several tissue classes can be smoothly incorporated into the classification framework by extending the original scheme by mixed tissue classes [3,4]. As classical methods of noise reduction are based on linear low-pass filtering, they necessarily blur the boundary between different tissues, leading to artificially created partial volume effects. Nonlinear

techniques based on anisotropic diffusion processes [5], which selectively stop the smoothing process at spatial positions with large intensity gradients, have been established during the past decade as effective tools for noise reduction, while preserving anatomical structures at tissue boundaries.

Several techniques have been developed for the correction of the spatial intensity bias resulting, for example, from the inhomogeneity of the RF field during MR image acquisition. One possibility considered is the implementation of bias-field correction as a preprocessing step, using generic assumptions about the expected distortions [6, 7]. As an alternative, expectation maximization has been proposed as an iterative framework to perform classification and bias correction simultaneously [8].

One important limitation of the above intensity-based classification framework is that it handles pixels in the image completely independently. This means that the result of the segmentation is invariant to the actual positions of the voxels in the image. This assumption is of course highly nonrealistic as intensities of nearby voxels are usually strongly correlated. This correlation between single pixels can explicitely be described by spatial interaction models. Spatial correlation between the single pixels can be introduced using more or less complex interaction models as, for example, Markov random fields [9, 10] and integrated into the classification framework. As an alternative, postprocessing techniques can be used to correct for erroneous classification. One popular technique is based on mathematical morphology [11], which allows the identification and correction of wrongly classified pixels based on morphological criteria, like the presence of very small, isolated tissue clusters [3]. The latter process is illustrated by the identification of the brain mask on a neuroradiological MR slice Fig. 14.3.

14.3 Knowledge-Based Automatic Segmentation

Even the most sophisticated pre- and postprocessing techniques cannot, however, overcome the inherent limitation of the basically intensity-based methods, namely the assumption that segmentation can be carried out solely based on information provided by the actual image. This assumption is fundamentally

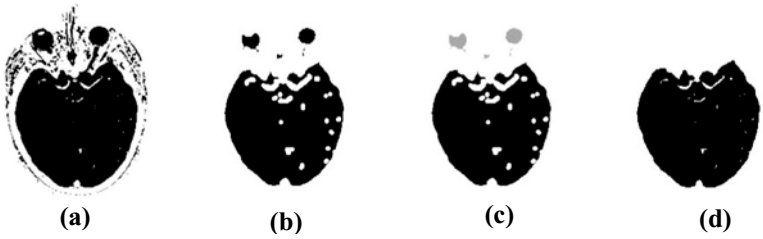


Figure 14.3: Brain segmentation based on morphological postprocessing. Image (a) shows the result of thresholding, which has been eroded (b) in order to break up unwanted connections between different organs. Brain tissue has been identified by connected component labeling (c) and has been dilated back to its original extent (d).

wrong, and the radiologist uses a broad range of related knowledge on the field of anatomy, pathology, physiology, and radiology in order to arrive at a reasonable image interpretation. The incorporation of such knowledge into the algorithms used is therefore indispensable for automatic image segmentation.

Different procedures have been proposed in the literature to approach the problem of representation and usage of prior knowledge for image analysis. Because of the enormous complexity of the necessary prior information, classical methods of artificial intelligence as the use of expert systems [12, 13] can offer only limited support to solve this problem.

14.3.1 Segmentation Based on Anatomical Models

During the past few years, the usage of deformable anatomical atlases has been extensively investigated as an appealing tool for the coding of prior anatomical information for image interpretation. The method is based on a representative deterministic [14] or probabilistic [15] image volume as an anatomical model. For this the actual patient data has to be spatially normalized, thus it has to be mapped onto the template that conforms to the standard anatomical space used by the model. The applied registration procedures range from simple parametric edge matching [16] and *rigid* registration methods over to increasingly more complex algorithms using *affine*, *projective*, and *curved* transformations. Other methods use complex physically inspired algorithms for elastic deformation or viscous fluid motion [17]. In the latter the transformations are constrained to

be consistent with the physical properties of deformable elastic solids or those of viscous fluids. Viscous fluid models are less constraining than linear elastic models and allow long-distance, nonlinear deformations of small subregions. In these formulations, the deformed configuration of the atlas is usually determined by driving the deformation using only pixel-by-pixel intensity similarity between the images if a reasonable level of automation has to be achieved. Common to all registration methods is the resulting dense vector field that defines the mapping of the subject's specific anatomy onto the anatomy template used for the atlas.

The usage of deformable atlases seems to be a very elegant way to use prior anatomical information in segmentation, as it allows to gain support from the success of current image registration research. Once the spatial mapping between the atlas and the individual data has been established, it can be used to transfer all spatially related information predefined on the atlas (as organ labels, functional information, etc.) to the actual patient image.

This approach is, however, fundamentally dependent on the anatomical and physiological validity of the generated mapping. It has to be understood, that a successful *warping* of one dataset into the other, does not guarantee that it also makes sense as an anatomical mapping. In other words, the fact that the registration result *looks* perfect offers no guarantee that it *makes sense* from the anatomical point of view. To warp a leg into a nose is perfectly possible, but will not allow any reasonable physiological interpretation.

To make the results of the registration sensible, i.e., useful for image segmentation, one has to solve the correspondence problem. This means that we have to ensure that the mapping establishes a correspondence between the atlas and the patient, which is physiologically and anatomically meaningful. For the time being, purely intensity driven registration cannot be expected to do so in general. Therefore, in the practice such correspondence usually has to be strongly supported using anatomical landmarks [18, 19]. Landmark identification needs, however, in most cases tedious manual work, compromising the quest for automatic procedures. The following section discusses one very popular way to address some of the mentioned fundamental problems of the atlas-based representation of anatomical knowledge. It can, however, hardly be expected that any of the individual methods alone can successfully deal with all aspects of automatic segmentation, and first attempts to combine different approaches have already been published [20].

14.3.2 Segmentation Based on Statistical Models

Anatomical structures show a natural variation for different individuals (*interindividual*) and also for the same individual (*intraindividual*) over time. Obvious examples for intraindividual variation of organ shape are the lungs or the heart that both show cyclic variation of their shape. In contrast the bladder shows noncyclic shape variations that mainly depend on its filling. Several researchers have proposed to model the natural (large, but still strongly limited) variability of *inter-* as well as *intraindividual* organ shapes in a statistical framework and to utilize these statistics for model-based segmentation. The idea is to code the variations of the selected shape parameters in an observed population (the training set) and characterize this in a possibly compact way. This approach overcomes the limitations of the basically static view of the anatomy provided by the altases from the preceding section.

Such methods fundamentally depend on the availability of parametric models suitable to describe biological shapes. These frameworks always consider variation of shape, but may also include other characteristics, such as the variation of intensity values. Several methods have been proposed for such parametric shape descriptions, as deformable superquadrics augmented with local deformation modeling [21, 22], series expansions [23, 24], or simply using the coordinates of selected organ surface points as used by Cootes in [25] for his point distribution models.

To model these statistics the *a priori* probability $p(\mathbf{p})$ of a parameter vector $\mathbf{p}$ and eventually the conditional probability $p(\mathbf{p} \mid \mathbf{D})$ under the condition of the sensed data D are estimated by a set of samples. Estimation of the probability $p(\mathbf{p})$, however, requires that the entities of the sample set are described in the same way. If, for example, the parameter vector $\mathbf{p}$ simply consists of voxel coordinates then it is essential, that the elements of the parameter vectors of the different entities always describe the position of the *same* point on the different entities at the same position in the vector. To find these corresponding points on the different realizations is an important prerequisite for the generation of statistical shape models. However, it proves difficult, especially as there in no real agreement on what correspondence exactly is, and how it can be measured.

Correspondence can be established in two ways, either discrete or continuously. In the discrete case the surfaces are represented as point sets and the correspondence is defined by assigning points in different point sets to each

other. In the continuous case, parameterizations of the surfaces are defined such that the same parameter values parameterize corresponding positions on the surface.

Looking at the discrete case the most obvious method is to define correspondences manually. To do so a number of landmarks need to be defined on each sample by the user. This method has been successfully used by [26]; however, this technique clearly requires extensive user input, making it only suitable when very limited number of points is regarded. Another possibility when dealing with discrete point sets offers the softassign Procrustes matching algorithm as described by [27]. The algorithm tackles the problem of finding correspondences in two point sets using the Procrustes distance shape similarity measure [18] that quantifies shape similarity.

The most common approach for the approximation of continuous correspondences in 2D is arc-length parameterization. Thus, points of the same parameter on different curves are then taken to be corresponding. This approach heavily depends on the availability of features and is thus bound to fail if the same features can not be located in both modalities. An other interesting view on continuous correspondences in 2D is given by [28], who defines correspondence between closed curves C_1 and C_2 as a subset of the product space $C_1 \times C_2$. Kotcheff and Taylor presents in [29] an algorithm for automatic generation of correspondences based on eigenmodes. In [30] Kelemen *et al.* shows a straight forward expansion of arc-length parameterization based correspondence of curves to establish correspondences on surfaces. They establish correspondence by describing surfaces of a training set using a shape description invariant under rotation and translation presented in [24].

Once the parameterization is selected, the anatomical objects of interest are fully described (at least from the point of view of the envisioned segmentation procedure) by the resulting parameter vector $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$, where n can of course be fairly large for complex shapes. Possible variations of the anatomy can be precisely characterized by the joint probability function of the shape parameters p_i, information of which can be integrated into a stochastic Bayesian segmentation framework as a prior utilizing the knowledge gained from the training data for constraining the image analysis process [22, 31]. It has to be, however, realized that the usually very limited number of examples in the training set forces us to very strongly limit the number of parameters involved in a fitting procedure. A very substantial reduction of the number of parameters can

be achieved based on the fact that the single components of the vector $\mathbf{p}$ are usually highly correlated. A simplified characterization of the probability density is possible based on the first- and second-order moments of the distribution (for a multivariate Gaussian distribution this description is exact). The resulting descriptors are

- the mean shape:

$$\overline{\mathbf{p}} = \frac{1}{N} \sum_{j=1}^{N} \mathbf{p}_j, \tag{14.1}$$

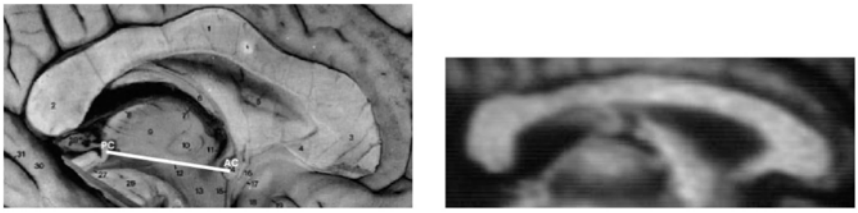
  where the training set consists of the N examples described by the parameter vectors $\mathbf{p}_j$;

- the covariance matrix of the components of the parameter vectors:

$$\Sigma = \frac{1}{N-1} \sum_{j=1}^{N} (\mathbf{p}_j - \overline{\mathbf{p}}) \cdot (\mathbf{p}_j - \overline{\mathbf{p}})^T \tag{14.2}$$

The existing correlations between the components of the vectors $\mathbf{p}$ can be removed by principal component analysis, providing the matrix $\mathbf{P_p}$ constructed from the eigenvectors of the covariance matrix Σ. Experience shows that even highly complex organs can well be characterized by the first few eigenvectors with the largest eigenvalues. This results in a description called active shape model [32], which allows to reasonably approximate the full variability of the anatomy by the deviation from the mean shape as a linear combination of a few eigenmodes of variation. The coefficients of this linear combination provide a very compact characterization of the possible organ shapes.

The automatic extraction of the outline of the corpus callosum on midsagittal MR images [33] nicely illustrates the basic ideas of using active shape models for segmentation. Figure 14.4 shows the region of interest covering the corpus callosum on a brain section (a) and on an MR image slice (b). Several examples have been hand-segmented, providing a training set of 71 outlines, which have been parameterized by Fourier coefficients up to degree 100. In order to incorporate not only shape-related but also positional variations into the statistical model, the contours have been normalized relative to a generally accepted neuroanatomical coordinate system, defined by the anterior and posterior commissures (Fig. 14.4). The training data used and the shape model resulting from the principle component analysis is illustrated in Fig. 14.5. As image (b) illustrates,

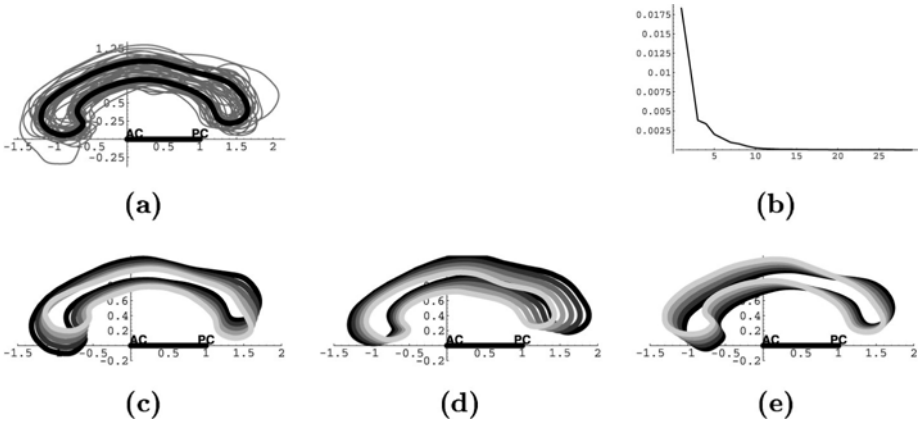


(a) **(b)**

Figure 14.4: (a) The corpus callosum from an anatomical atlas and (b) the corresponding region of interest in a midsagittal MR image. On the left image the connecting line between the anterior commissure (AC) and the posterior commissure (PC), which is used for normalization, is also shown.

the largest 12 eigenvalues (defined by the 400 original parameters) already reasonably represent the variability (covering about 95% of the full variance).

This statistical description can easily be used as a parametric deformable model allowing the fully automatic segmentation of previously unseen images (apart from the definition of the stereotactic reference system). Based on the



(a) **(b)**

(c) **(d)** **(e)**

Figure 14.5: Building the active shape model for the corpus callosum. (a) The 71 outlines of the training set normalized in the anatomical coordinate system defined by the anterior and posterior commissures (AC/PC). The eigenvalues resulting from the principal component analysis are plotted in (b), while the eigenvectors corresponding to the three largest eigenvalues are illustrated in (c), (d), and (e). The deformations which correspond the eigenmodes cover the range $-\sqrt{2}\lambda_k$ (light gray) to $+\sqrt{2}\lambda_k$ (dark gray).

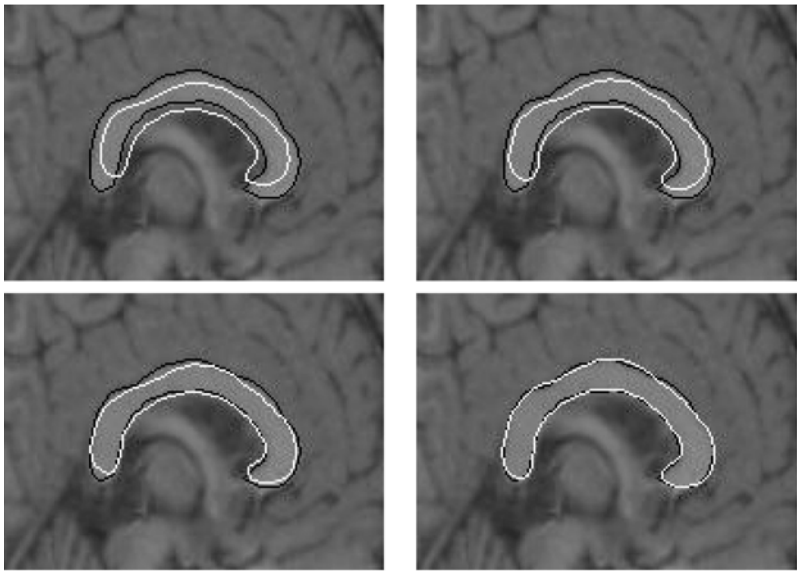
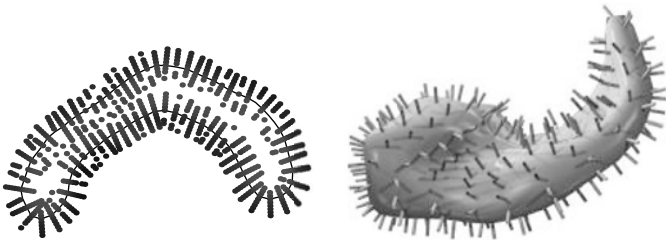


Figure 14.6: Segmentation of the corpus callosum. The top-left image shows the initialization, resulting from the average model and a subsequent match in the subspace of the largest four deformation modes. The other images (top right, lower left, and lower right) illustrate the deformation of this model during optimization using all selected deformation modes, allowing fine adjustments. The black contour is the result of a manual expert segmentation.

concept of deformable contour models or snakes [34] (see section 14.4.3), the corpus callosum outline can be searched in the subspace spanned by the selected number of largest eigenmodes using the usual energy minimization scheme as illustrated in Fig. 14.6. The efficiency of the fit can be vastly increased by incorporating information about the actual appearance of the organ on the radiological image, for example, in the form of intensity profiles along its boundary, as illustrated in Fig. 14.7(a), leading ultimately to the usage of integrated active appearance models [35] incorporating the shape and gray-level appearance of the anatomy in a coherent manner.

The illustrated ideas generalize conceptually very well to three dimensions, as illustrated on the anatomical model of the basal ganglia of the human brain shown in Fig. 14.8. The corresponding active shape model has been successfully applied for the segmentation of neuroradiological MR volumes [36]. Remaining interactions needed for the establishment of the anatomical coordinate system



(a) (b)

Figure 14.7: Intensity profiles along the boundary of a (a) 2-D and a (b) 3-D object.

can be eliminated using automated adaptation of the stereotactical coordinate system [37].

The approach presented in this chapter can be extended to multiorgan matching, since they are spatially related to each other. The prostate, for example, is placed at an inferior dorsal position to the bladder. The bladder broadly changes its shape due to its filling and this deformation also influences the shape and position of the prostate. This correlation of the position and shape of the organs can be modeled by incorporating multiple organs in the shape statistics. To do so the coefficient vectors $\mathbf{p}_i$ of the n incorporated organs can be gathered in one single coefficient vector $\mathbf{p}_{\text{comb}} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n)$. As modeling the relative position of the organs is believed to be one of the major benefits of multiorgan modeling, the centers of gravity must be considered in the statistics. In

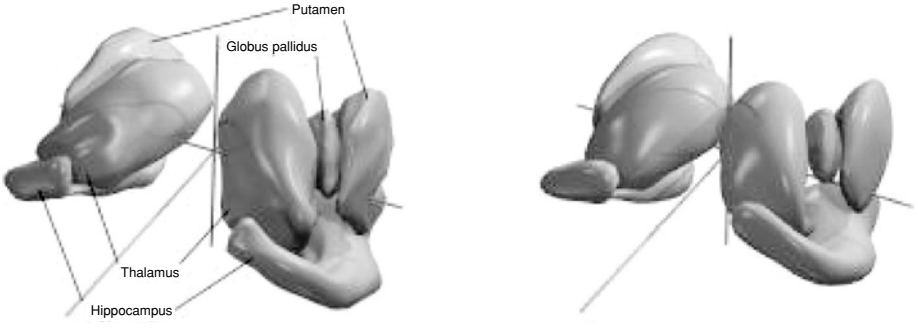
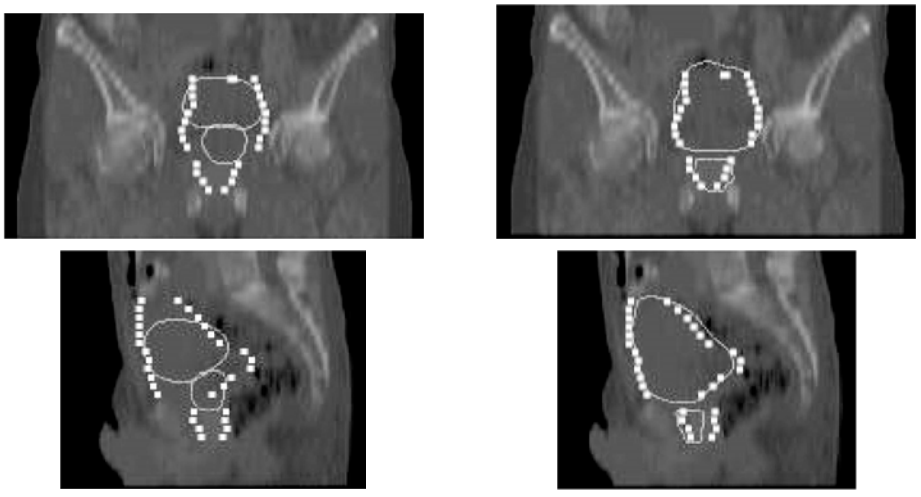


Figure 14.8: Three-dimensional model of the basal ganglia of the human brain. On the left an individual anatomy from the training set is shown, while the the average model is presented in the right image.



Mean model Correct eignmodes for bladder

Figure 14.9: Prediction of the position of the prostate for a known bladder shape.

particular the relative positions with respect to the origin of a common coordinate system of the combined organs must be modeled. There are different possibilities to choose this reference coordinate system. One possible and intuitive choice for a reference coordinate system would be the center of gravity of one of the organs.

Figure 14.9 shows that the position of the prostate depending on the shape of the bladder is modeled reliably. Here, the mean bladder–prostate model is shown on the left. In the right the first 10 eigenmodes were added, so that they best approximate the bladder. As can be seen in the figure, the position of the prostate is also approximately found, although no information on the prostate was included.

It should be noted that the establishment of correspondence is still a major matter of concern while the training set is created, which further complicates the generation of suitable data collections for training. The intensive manual work needed is, however, *limited to the training phase*, while the actual segmentation of the unseen data is fully automatic. The correspondences including the spatial variability and radiological appearance of the anatomical landmarks are integrated into the statistical model and will be transfered to the new images during the fitting process.

14.4 Interactive Segmentation

In spite of the considerable success of knowledge-based automatic segmentation, generic algorithms capable to analyze and understand complex anatomical scenes cannot be expected to be available in the near future. The major reason for the slow progress is that current methods can cover only a very limited fragment of the whole spectrum of prior knowledge, which clinicians use when analyzing radiological images. Accordingly, available solutions can be applied only on very limited problem domains and individual solutions calling for major investment in research and development have to be sought for when trying to address different clinical questions.

However, the growing demand for computer support in the clinical praxis and the vast amount of data produced by routine radiological acquisitions urgently call for methods, which are capable to effectively deal with a broad spectrum of clinical problems without demanding unacceptable workload from the clinician. The key to such solutions is an optimal cooperation between the computer and the human operator, which allows the user to rely on the advantages of computerized image analysis (like reproducibility and stamina), while enabling him to contribute to the solution with the full scale of his domain knowledge. During the past 15 years a new family of segmentation algorithms emerged following this paradigm, which is coined as interactive segmentation (in contrary to fully manual methods, where the computer is simply used as a more or less intelligent drawing tool).

This section presents two classes of interactive segmentation techniques, namely, graph-based methods and physically based deformable models. The Live-Wire paradigm has been chosen as a representative of the first class, while Snakes will serve as an example of the second. In addition, two recent extensions of the classical snake's definition will be discussed, namely, Ziplock Snakes and Tamed Snakes. Finally, the extension of the snakes approach into three dimensions will be discussed. This review is by no means exhaustive, but intended to give a brief introduction into a wide area of different interactive segmentation methods that have been proposed during the last decades. The description of two different segmentation prototypes should nevertheless enable the reader to understand the related key components and challenges.

14.4.1 Live-Wire Segmentation

The first class of algorithms reviewed, which are usually referred to as Live-Wire algorithms [38], belongs to the field of dynamic programming. The foundations of these algorithms lie in the F* algorithm [39], and will briefly be sketched here.

In a Live-Wire algorithm, the image I is considered as a discrete neighborhood graph, where each pixel corresponds to a node in the graph. Generally an 8-Neighborhood ($\mathcal{N}_8$, Moore neighborhood) is defined, so that diagonal connections are allowed. A cost function $C(I)$ assigns a value to each node in the graph. Typically, the cost function is based on local features, for example, the output of an edge detector. Once a suitable cost function is defined, the segmentation task is reduced to a minimal cost path search problem between two points in the image graph. These points are usually selected by the user, who defines a starting point s and then drags the endpoint e with the mouse toward a desired location. During dragging the algorithm repeatedly evaluates the minimal cost path $P_{\min}$ from the starting point s to the current location e as illustrated in Fig. 14.10. In order to evaluate the minimal cost path $P_{\min}$, a path array P has to be constructed that accumulates the total costs from the starting point s to any point of the image.

All values of the path array P are initially set to infinity except the start vertex s, which is assigned a value equal to its cost $C(s)$. The elements of the

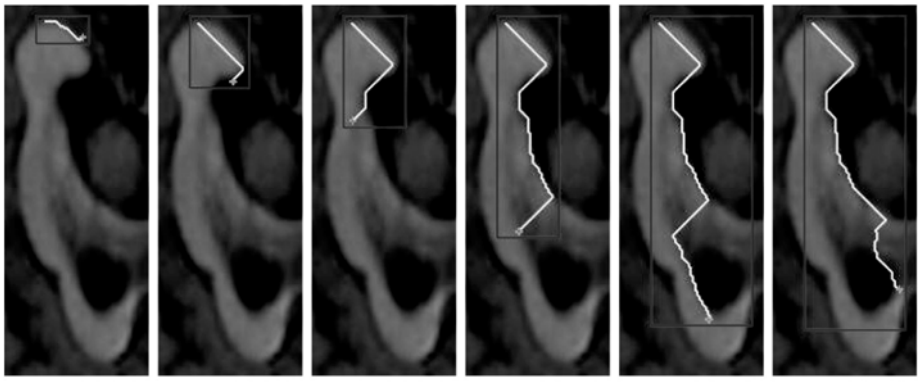


Figure 14.10: Segmentation of the corpus callosum using the F* algorithm. Note the jumping behavior between the last two images due to global optimum computation.

path array are then updated iteratively until convergence, i.e., until no more values in the array are changed in one iteration. The values of the ith row of P are first adjusted from left to right by the following rule

$$P(i,j) = \min \begin{pmatrix} P(i-1,j-1) + C(i,j) \\ P(i-1,j) \quad\quad + C(i,j) \\ P(i-1,j+1) + C(i,j) \\ P(i,j-1) \quad\quad + C(i,j) \\ P(i,j) \end{pmatrix}, \tag{14.3}$$

and then from right to left according to

$$P(i,j) = \min\left(P(i,j+1) + C(i,j), P(i,j)\right). \tag{14.4}$$

Additional iterations alternate between a bottom-to-top pass (with reversed indices, so that the bottom row corresponds to $i = 1$) followed by a top-to-bottom pass. In each pass the updating rules 14.3 and 14.4 are applied. Once the minimal cost array P has been generated, the minimal cost path $P_{\min}$ from any point e' to the starting point s can easily be computed by backtracking from e' to s without recalculating P, thus making the algorithm very fast.

As can be seen in Fig. 14.10, the generated segmentation line approximates the edges present between the start- and endpoint selected. The continuous computation of a global optimum leads to the jumping behavior of the algorithm, as can be observed in the last two images of Fig. 14.10. The last segment of the segmentation line abruptly changes the edges that are approximated.

Modifications of the resulting segmentation cannot be directly integrated in the concept of Live-Wires as the path array P is fixed solely based on the image on initialization, rendering any post processing a cumbersome task. To tackle the discontinuities, the user can generally fix the endpoint of the segmentation line so far if estimated adequately, thus freezing the wire and starting a new segment. A piecewise construction of the segmentation is obviously mandatory for any closed contour.

Different extensions have been proposed to improve the behavior of the Live-Wire algorithms. These extensions include, but are not limited to the definition of more complex cost functions [38], advanced edge feature detections [40] and the extension to 3D [41–43]. The main advantages of the Live-Wire approach are the relatively simple implementation and the computational speed, as the

shortest path can be computed online while the user drags the mouse, thus providing direct feedback.

In contrast to Snakes, which will be described next, the selected path is a global optimum of the cost function defined over the complete image, whereas Snakes iteratively adapt their contours based on local information, which will very likely represent a local optimum of their respective target function. Global optimum is a desirable mathematical property in optimization; it is, however, not obvious that the best segmentation is equivalent to this definition of an optimum.

14.4.2 Snakes

The second class of algorithms presented intends to overcome some of the limitations of the graph-based approaches. The former allows the segmentation line, respectively surface, to have individual properties that are not related to the image, but rather to physical properties of some material. The segmentation process is no longer solely based on the image, but regularized by the constraineds imposed by the physical model. This model introduces some generic knowledge of general organ's shape encoded in the elasticity of the material. The algorithms of this section belong to the field of physically based deformable models. In this section, the basics of Snakes are first depicted, followed by the presentation of two extensions that have been proposed during the last few years.

Traditional Snakes used for image segmentation are polygonal curves to which an objective function is associated. This function combines an "image term," which measures either edge strength or edge proximity, and a regularization term, which accounts for tension and curvature. The curve is deformed in order to optimize a score and, as a result, to match the image edges. The optimization typically is global, i.e., it takes edge information into account along the whole curve simultaneously, but only considers local image information, i.e. image intensities close to the curve.

Snakes were originally introduced by Kass *et al.* [44] and are modeled as time-dependent 2-D curves defined parametrically as

$$\mathbf{v}(s,t) = (x(s,t),\, y(s,t))_{0\le s\le 1}, \tag{14.5}$$

where s is proportional to the arc length, t the current time, and x and y the curve's image coordinates. The Snake deforms itself as time progresses so as to

minimize an energy term that combines image, internal, and external energies:

$$E = E_{\text{image}} + E_{\text{int}} + E_{\text{ext}}. \tag{14.6}$$

These energies are derived by integration along the curve. The forces resulting from the minimization of the image energy E_{image} guide the Snake to match the desired image features. This image energy is derived by integrating over a potential $P(\vec{v}(s,t))$ from an image feature map, i.e.

$$E_{\text{image}}(\mathbf{v}) = -\int_0^1 P(\mathbf{v}(s,t))\,\mathrm{d}s \tag{14.7}$$

A typical choice is to take $P(\mathbf{v}(s,t))$ to be equal to the magnitude of the image gradient, that is

$$P(\mathbf{v}(s,t)) = |\nabla I(\mathbf{v}(s,t))|\ , \tag{14.8}$$

where I is either the image itself or the image convolved by a Gaussian kernel. As for the Live-Wire cost function, many different feature maps have been suggested in the past [45–47], yet the results are comparable.

The internal energy term E_{int} models the physical behavior of the material describing the Snake. Using the elastic rod model, E_{int} is taken to be

$$E_{\text{int}}(\mathbf{v}) = \frac{1}{2}\int_0^1 \alpha(s)\left|\frac{\partial \mathbf{v}(s,t)}{\partial s}\right|^2 + \beta(s)\left|\frac{\partial^2 \mathbf{v}(s,t)}{\partial s^2}\right|^2 \mathrm{d}s. \tag{14.9}$$

The parameters $\alpha(s)$ and $\beta(s)$ are arbitrary functions that regulate the curve's tension and rigidity and are commonly supplied by the user. It has been shown that they can be chosen in a fairly image-independent way [46]. Generally $\alpha(s)$ and $\beta(s)$ are defined as constant along the curve, i.e. $\alpha(s) = \alpha$ and $\beta(s) = \beta$. Other authors have proposed to dynamically adjust the values of α and β along the curve by a feed-back strategy [48].

The segmentation process performed with Snakes is governed by the minimization of the term $\int E(\mathbf{v})\,\mathrm{d}t$. This amounts to using Hamilton's principle in variational calculus to derive the Euler–Lagrange equations of motion. The resulting equation of motion for the basic Snake described here can be written as

$$-\alpha\, v_{ss} + \beta\, v_{ssss} = -P_v, \tag{14.10}$$

using subscripts to denote derivatives. For computational purposes the equation has to be discretized. The model $\mathbf{v}$ is taken to be a polygonal curve defined by a set of vertices $\mathbf{v}^{[t]} = (x_i^{[t]}, y_i^{[t]})_{0 \le i \le n-1}$. It is customary to use central differences, yielding the discrete counterpart to Eq. (14.10)

$$-\alpha \left(v_{i-1}^{[t]} - 2v_i^{[t]} + v_{i+1}^{[t]}\right) + \beta \left(v_{i-2}^{[t]} - 4v_{i-1}^{[t]} + 6v_i^{[t]} - 4v_{i+1}^{[t]} + v_{i+2}^{[t]}\right) = -P_v\big|_{\mathbf{v}^{[t-1]}}, \tag{14.11}$$

which has to be solved for every single vertex $v_i^{[t]}$ simultaneously. Using matrix notation, the linear system of equations can be written as

$$\mathbf{K}\mathbf{v}^{[t]} = -P_v\big|_{\mathbf{v}^{[t-1]}}, \tag{14.12}$$

where $\mathbf{v}$ stands for either $(x_0, \ldots, x_{n-1})$ or $(y_0, \ldots, y_{n-1})$. The stiffness matrix $\mathbf{K}$ is pentadiagonal and singular, thus Eq. (14.12) cannot be solved by direct inversion of $\mathbf{K}$. To be able to solve the Snake Eq. (14.10), two different methods will be described. First an iterative solution is presented, which stems from the original Snakes framework. Ziplock Snakes, which will be introduced in section 14.5.1.1, incorporate boundary conditions into Eq. (14.12), so that the equation can be solved by inversion of $\mathbf{K}$. In the original approach, additional terms are incorporated into Eq. (14.10) that introduce a temporal development of the Snake.

14.4.2.1 Dynamic Terms

Dynamic terms have been introduced to account for kinetic energy and velocity-dependent friction, leading to a more physically reasonable movement of the Snake [49]. The kinetic energy term $E_{\mathrm{K}}(\mathbf{v})$ is set to

$$E_{\mathrm{K}}(\mathbf{v}) = \frac{1}{2} \int_0^1 \mu(s) \left| \frac{\partial \mathbf{v}(s,t)}{\partial t} \right|^2 \mathrm{d}s, \tag{14.13}$$

where $\mu(s)$ represents the mass of the Snake.

14.4.2.2 Energy Dissipation

The physical system described so far is energy conserving. The introduction of energy dissipation results in a more realistic physical behavior and can be

modeled using a Rayleigh dissipation functional

$$D(\mathbf{v}_t) = \frac{1}{2} \int_0^1 \gamma(s) \, |\mathbf{v}_t|^2 \, \mathrm{d}s, \tag{14.14}$$

$\gamma(s)$ being the damping coefficient.

The segmentation process is now determined by the minimization of the term $\int E(\mathbf{v}) + E_{\mathrm{K}}(\mathbf{v}) + D(\mathbf{v}_t) \, \mathrm{d}t$. Using the Hamiltonian principle the following Euler–Lagrange differential equation results

$$\mu \, v_{tt} + \gamma \, v_t - \alpha \, v_{ss} + \beta \, v_{ssss} = -P_v, \tag{14.15}$$

where $\mu(s)$ and $\gamma(s)$ are considered constants along the curve. Forward differences are used to approximate the time derivatives, resulting in a linear system of equations which can be formulated in matrix notation as

$$((\mu + \gamma)\mathbf{I} + \mathbf{K}) \cdot \mathbf{v}^{[t]} = -P_v\big|_{\mathbf{v}^{[t-1]}} + (2\mu + \gamma) \, \mathbf{v}^{[t-1]} - \mu \, \mathbf{v}^{[t-2]}. \tag{14.16}$$

The role of the damping becomes evident, as the condition of the stiffness matrix $\mathbf{K}$ is improved through the damping term $(\mu + \gamma)\mathbf{I}$. This term has to be selected in a reasonable manner to prevent the Snake to be "glued," which would be the case for $|(\mu + \gamma)\mathbf{I}| \gg |\mathbf{K}|$. With appropriate selections for μ and γ, a well-conditioned linear system results, so that the term $(\mu + \gamma)\mathbf{I} + \mathbf{K}$ can be inverted and Eq. (14.16) solved for $\mathbf{v}^{[t]}$ yielding an iterative algorithm to solve the Snake equation of motion.

The Lagrangian formalism allows to unify forces from very different type of sources. The target function is extended to incorporate more energy terms so that the final energy to be minimized is of the form $E_{\mathrm{tot}} = \sum_i E_i$. Some basic extensions that have been presented are summarized.

14.4.2.3 External Forces

User interaction is introduced through external forces, so that the Snake can be modified manually [49]. Two type of forces are commonly used to attract or repulse the Snake from the current mouse position. Attraction can be modeled by introducing a virtual spring connecting the mouse position $\mathbf{m}$ with a point $\mathbf{p}$ on the Snake, that adds a term $k(\mathbf{p} - \mathbf{m})^2$ to the external energy E_{ext}, where the spring constant k is a parameter. To push the Snake away from an undesired local energy minimum, a "volcano" is introduced by adding an energy function

proportional to $\frac{1}{r^2}$ to the external energy, where r is the distance of a point from the volcano center. Obviously, special care is required to avoid instabilities near $r = 0$.

14.4.2.4 Balloon Forces

Traditional Snakes exhibit a tendency to shrink, as they reach an absolute minimum when contracted into a single point under a flat potential field, i.e., a homogeneous image. While the correction of the energy term E_{int} to enforce parameterization according to the arc-length can prevent this behavior, the resulting governing equations are highly nonlinear. Instead, an inflating force can be introduced, which expands a closed Snake like a balloon [50]. Denoting the unit vector normal to the curve at point $\mathbf{v}(s)$ with $\mathbf{n}(s)$, the additional energy term becomes

$$E_{\text{B}}(\mathbf{v}) = \int_0^1 \mathbf{n}(s)\,\mathrm{d}s. \tag{14.17}$$

14.4.2.5 Gravitational Forces

Similar to the balloon forces, a gravitational force can be defined [51]. Interpreting the image intensity as the z-dimension, the energy term is defined as

$$E_{\text{G}}(\mathbf{v}) = \int_0^1 \mathbf{g}(s)\,\mathrm{d}s, \tag{14.18}$$

where $\mathbf{g}(s)$ is the gravitation vector $(0, 0, -g)$. The Snake is then accelerated in negative z-direction, so that the model seems to "falls" on an object.

14.4.3 Deformable Surface Models

The basic concepts of Snakes—minimization of an energy term through optimization—can easily be generalized to three dimensions. Additional effort is required only to handle the parameterization problem adherent to 2-D manifolds.

In contrast to 2-D active contour models, where arc length provides a natural parameterization, 2-D manifolds as used for 3-D deformable models pose a complex, topology, and shape-dependent parameterization problem. Parameterizing

a surface effectively is difficult because there is no easy way to distribute the grid points evenly across the surface.

A generalized deformable surface model is defined as

$$\mathbf{v}(\omega, t) = (x(\omega, t), y(\omega, t), z(\omega, t)) \tag{14.19}$$

where $\omega \in \Omega \subset \Re^2$ is a suitable parameterization, t the current time, and x, y and z are the corresponding coordinate functions of the surface. Analogous to the 2D case, the surface deforms itself so as to minimize its image potential energy. Instead of the elastic rod model, the thin plate under tension model is employed to regulate the model's shape during energy minimization. Thus, the term E_{int} has the following form:

$$E_{\text{int}}(\mathbf{v}) = \iint_{\Omega} \tau(\omega) \left[\left| \frac{\partial \mathbf{v}}{\partial \phi} \right|^2 + \left| \frac{\partial \mathbf{v}}{\partial \theta} \right|^2 \right] + \rho(\omega) \left[\left| \frac{\partial^2 \mathbf{v}}{\partial \phi^2} \right|^2 + 2 \left| \frac{\partial^2 \mathbf{v}}{\partial \phi \partial \theta} \right|^2 + \left| \frac{\partial^2 \mathbf{v}}{\partial \theta^2} \right|^2 \right] \mathrm{d}\omega \tag{14.20}$$

where $\rho(\omega) = 1 - \tau(\omega)$ for convenience and $\omega = (\phi, \theta)$. The surface tension parameter τ is a user-supplied parameter in the range 0..1, varying the behavior of the surface between a thin plate ($\tau = 0$) and a membrane ($\tau = 1$). When endowing the surface with a mass μ and embedding it into a viscous medium the corresponding energy terms are the same as for the traditional Snake:

$$E_{\text{K}}(\mathbf{v}) = \frac{\mu}{2} \iint_{\Omega} \left| \frac{\partial \mathbf{v}(\omega, t)}{\partial t} \right|^2 \mathrm{d}\omega \qquad D(\mathbf{v}_t) = \frac{\gamma}{2} \iint_{\Omega} |\mathbf{v}_t|^2 \, \mathrm{d}\omega \tag{14.21}$$

The Euler–Lagrange differential equation for constant parameters μ, γ, τ can be formulated

$$\mu \, \mathbf{v}_{tt} + \gamma \, \mathbf{v}_t - \tau \, \Delta \mathbf{v} + (1 - \tau) \Delta^2 \mathbf{v} = -\frac{\partial P}{\partial \mathbf{v}}, \tag{14.22}$$

where $\mathbf{v}$ stands for either x, y, or z. These coupled differential equations can be solved numerically as in two dimensions.

14.4.4 Tamed Snakes

Despite their ability to approximate objects with little user input, the approaches reported so far lack an intuitive manipulation semantic. The primitive manipulation metaphors presented, spring and volcano, can only be applied directly on the Snake's curve $\mathbf{v}$. It would be desirable though to have a more powerful interaction method at hand, for example, to modify parts of the shape on a coarse

scale, while keeping small details on a finer scale intact. Tamed Snakes combine the hierarchical modeling and Snake-like edge delineation. They adhere to the concept of hierarchical shape representations with several scales of resolution to provide the necessary interactive modeling capabilities while being suitable for numerical simulations.

Hierarchical modeling consists of (a) an iterative refinement of the geometry, which defines a hierarchy of representations and (b) a local detail encoding, which represents the details on a finer level with respect to the next coarser one. Subdivision curves are best suited for such hierarchical modeling, as their representation implicitly comprises a hierarchy of refined shapes. These curves are constructed using univariate subdivision schemes defined as the iterative application of an operator that maps a given polygon $\mathcal{P}_l = [\mathbf{v}_i^{(l)}]$ to a refined polygon $\mathcal{P}_{l+1} = [\mathbf{v}_i^{(l+1)}]$, where l denotes the level of the hierarchy. Such an operator is given by two rules for computing the new so-called even vertices $\mathcal{P}_{l+1}^{\sharp} = \{\mathbf{v}_{2i}^{(l+1)}\}$ and the new odd vertices $\mathcal{P}_{l+1}^{\sharp} = \{\mathbf{v}_{2i+1}^{(l+1)}\}$.

The Tamed Snakes as introduced by Hug [52] employ the DLG-subdivision scheme [53], given by

$$\mathbf{v}_{2i}^{(l+1)} = \mathbf{v}_i^{(l)} \qquad \mathbf{v}_{2i+1}^{(l+1)} = \left(\frac{1}{2} + \omega\right)\left(\mathbf{v}_i^{(l)} + \mathbf{v}_{i+1}^{(l)}\right) - \omega\left(\mathbf{v}_{i-1}^{(l)} + \mathbf{v}_{i+2}^{(l)}\right). \quad (14.23)$$

As the even vertices remain unchanged the subdivision operator has interpolating behavior. The free tension parameter ω has to be chosen inside the interval $0 < \omega < \frac{1}{8}$ to obtain a limit curve that has a continuous tangent vector.

Local details, i.e. transformations of the vertices $\mathbf{v}_i^{(l)}$ from their given position, are encoded by establishing a local coordinate system $f_i^{(l)}$ in each vertex $\mathbf{v}_i^{(l)}$ and by representing the details with respect to this frame.

The subdivision scheme suggests to start the segmentation process with a reasonably coarse model and to iteratively adjust and refine the control vertices of the resulting curve. Since the subdivision scheme is interpolating, only the odd vertices of the current level must be adjusted to align with a correct boundary position before proceeding to the next finer level. In doing so, the prediction of the refinement operator improves continuously with respect to the vertex position on the next finer level and converges to the correct boundary position.

The traditional Snake energy has to be modified to combine the described coarse-to-fine approach with the Snake-like edge tracking. Tamed Snakes replace the elastic rod model term (Eq. 14.9) by a spring energy similar to the

external energy term introduced earlier for mouse interaction with Snakes. The springs are attached to each odd vertex $\mathbf{v}_i \in \mathcal{P}_l^\sharp$, so that the vertices $\mathbf{v}_i$ snap to the correct boundary position within the vicinity of their starting positions $\mathbf{v}_i(0) = \mathbf{v}_i \mid_{t=0}$. Assuming a good initialization, the imposed restriction on the search space to the local neighborhood is reasonable. The spring constant $k^{(l)}$ can be increased with each subdivision step to further restrict the search area, as the error of the subdivision operator's prediction tends to decrease.

Besides the spring energy, Tamed Snakes incorporate a kinetic energy E_K, an image energy E_image and a Rayleigh dissipation functional $D(\mathbf{v}_t)$. The resulting Euler–Lagrange equation of motion for Tamed Snakes describing the motion of all odd mass points $\mathbf{v}_i$ at time t is

$$\forall \mathbf{v}_i \in \mathcal{P}_l^\sharp : \mu_i \frac{\partial^2 \mathbf{v}_i(t)}{\partial t^2} + \gamma_i \frac{\partial \mathbf{v}_i(t)}{\partial t} + k_i^{(l)}(\mathbf{v}_i(t) - \mathbf{v}_i(0)) = -\nabla_\perp P(I)\big|_{\mathbf{v}_i(t)} \qquad (14.24)$$

In order to prevent the control vertices from drifting along the boundary, the gradient of the potential is projected onto the normal direction of the curve, denoted by $\nabla_\perp$ in the previous equation.

The segmentation process using Tamed Snakes is depicted in Fig. 14.11. The initialization has a strong impact on the additional manual editing required in finer levels. For the presented case, user interaction was required on a few points in the first and second subdivision level. Because of the limited number of vertices in these levels and the ability to better predict new positions at finer levels, Tamed Snakes proof themselves to alleviate the interactive segmentation

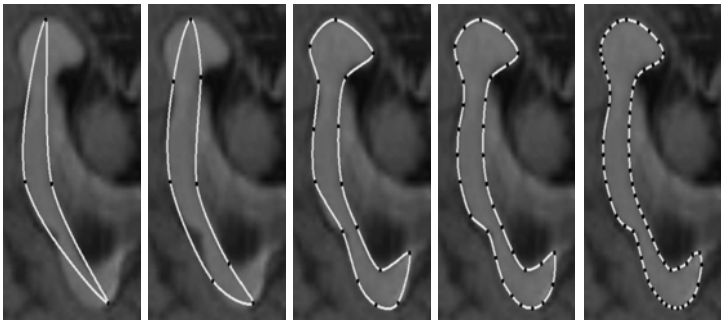


Figure 14.11: Segmentation using Tamed Snakes. The adaptive character of the subdivision scheme is clearly visible. User interaction was required in the second and third image.

task. In case of clear boundaries though, the segmentation is not as fast and elegant as with traditional Snakes.

14.4.5 Tamed Surfaces

Analogous to the 2-D case, a subdivision scheme is employed to generate a hierarchy of triangles, which is subject to the governing equations of the physical model. The modified Butterfly subdivision scheme for triangulated surfaces has been suggested for this purpose [54]. It was originally introduced by the same authors [55] as the DLG subdivision scheme in two dimensions and exhibits similar properties: it has interpolating behavior and has a tangent-continuous limit surface, i.e., C^1. As for most subdivision methods for triangulated surfaces, quaternary subdivision is used. To correct for degeneracies resulting from topologically irregular neighborhoods, i.e. for vertices with valence other then six, the extensions proposed by Zorin [56] have to be incorporated, hence the term "modified" in the name of the scheme.

Considering the extensions, the weights for the new vertices $\mathcal{V}_l^{\sharp}$ are computed as a function of the valence of the vertices $\mathbf{v}_i$. To solve Eq. (14.22) on triangular meshes, the Laplace operator has to be replaced by a discrete operator $\mathcal{L}$. One example of such an operator is the so-called umbrella-operator $\mathcal{U}$ introduced by Kobbelt [57]:

$$\mathcal{U}(\mathbf{v}_i) = \frac{1}{n_i} \sum_{j \in \mathcal{N}_1(i)} \mathbf{u}_j - \mathbf{v}_i \tag{14.25}$$

where n is the valence of the vertex $\mathbf{v}_i$. This operator clearly does not consider the geometric constellation of the neighborhood of $\mathbf{v}_i$, but results in a simple computation of the Laplace operator with reasonable accuracy for regular meshes.

The approximation of differential operators on arbitrary, discrete 2-D manifolds poses a complex problem. In contrast to the 1-D situation, where the adjacent vertices are always the left and right neighbors of the current vertex, there exists no such fixed relationship for 2-D manifolds. Many different methods for the computation of discrete operators have been proposed in the past few years [58, 59].

At this point it has to be noted that practical implementations of 3-D snakes pose additional challenges that have to be considered. In general, the 3-D

situation requires a stronger shape regularization in order to preserve a valid mesh structure. The projection of the image forces into normal directions, as suggested in Eq. (14.24) can only be applied on finer levels, as the normals of the coarse mesh may point in rather odd directions. In the context of tamed models, fixing all even vertices of the coarser mesh $\mathcal{V}_l^\natural$ can have an adverse effect on the optimization leading to strong parametric distortions. Hug [54], therefore, recommends to "freeze" as few vertices as possible, depending on the quality of the underlying data.

14.5 Deformable Model Initialization

14.5.1 Background

An essential prerequisite of interactive segmentation, which affects overall accuracy as well as efficiency of the method, is the sound initialization of the underlying model. On the one hand, the initial guess has a critical impact on the quality of the segmentation outcome. On the other hand, tedious and time-consuming manual initialization procedures forfeit possible time savings of the segmentation phase.

Although these are well-known facts, emphasis in the literature is usually placed on extensions of the deformable models, while an initial position relatively close to the desired solution is assumed. Nevertheless, the determination of such an initial guess with mouse-based interfaces, especially in 3D, poses a problem.

In the following, two approaches to aid a user in the fast generation of an initialization for a deformable model are described. In the first method, *a priori* shape knowledge is used for efficient initialization, thus reducing the amount of user interaction. In the second approach, the human–computer interface itself is enhanced by using multimodal interaction metaphors stemming from virtual reality techniques.

14.5.1.1 Ziplock Snakes

Ziplock Snakes emphasize on the improvement of the result based on the user's initialization [60]. They reduce the requirements on the initialization while

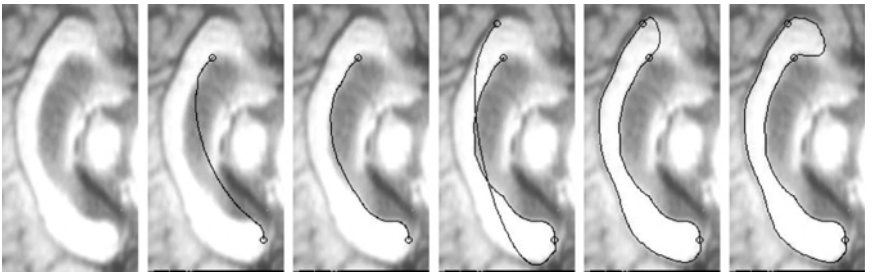


Figure 14.12: Segmentation process using Ziplock Snakes. It can be observed how the single segments are optimized from the user selected endpoints towards the center of the segments.

increasing the influence of this information. Traditional Snakes rely on the "closeness" of the initial Snake to the desired result. Depending on the underlying image, the term "closeness" transfers to an almost complete, manual delineation of the desired object's boundary. Ziplock Snakes in contrast require only the specification of the endpoints of the Snakes in the vicinity of clearly visible edge segments, which implies a well-defined edge direction. The system then optimizes the location of the user-supplied points to ensure that they are indeed good edge points, and extracts the associated edge directions. These anchor elements are used as boundary conditions and the edge information is then propagated along the Snake starting from them. The resulting behavior is visually similar to closing a zip, as can be observed in Fig. 14.12. The optimization of the energy term starts by defining the initial Snake as the solution of the corresponding homogeneous version of the system of differential Eq. (14.10). The selected endpoints provide the necessary boundary conditions $v(0)$, $v'(0)$, $v(1)$, and $v'(1)$ to solve this equation directly, i.e. Eq. (14.10) has a unique solution. At this stage the Snake "feels" absolutely no external image forces, as $-P_v = 0$ for the homogeneous case. Assuming that the user selects both endpoints near dominant edge fragments in the image, this initialization ensures that the Snake already lies close to its optimal position at both ends. During the ongoing iterative optimization process, the image potential P is turned on progressively for all the Snake vertices, starting from the extremities. Two types of Snake nodes are discerned, depending on whether the potential force field F_{P} is turned on (*active* nodes) for that vertex or not (*passive* nodes).

The user interaction closely resembles the Life-Wire approach: start- and endpoints of single segments have to be specified and the complete contour is assembled from several segments. The potential discontinuities arising at the connecting vertices are compensated by the fact that these vertices were selected on salient edges with clear directions.

Ziplock Snakes improve the overall convergence properties of Snakes and the probability of getting trapped in an undesirable local minimum is considerably reduced in most cases. However, gaps in object boundaries, misleading edges, and object outlines with low contrast represent insuperable obstacles that are quite usual in medical imagery.

14.5.1.2 Velcro Surfaces

The 3-D analogs of Ziplock Snakes are called Velcro surfaces, as their behavior mimics a piece of Velcro that is progressively clamped onto the surface of interest.

Following a natural extension from 1- to 2-D manifolds, points become lines. In the case of the Snakes under scrutiny this observation states that the initialization of 3-D models requires the specification of lines as boundary conditions. This conclusion comprises the original goal of the Ziplock framework—to reduce the user interaction. From the end-user's perspective, the specification of point landmarks for the initialization of the surface models is more desirable as it can be provided faster and more reliably. Velcro surfaces aim at such a landmark based initialization.

Assuming a set of anchor points and surface normals are given, a solution for the homogeneous equation (thin plate problem without external forces, $\tau = 0$, see Eq. (14.22))

$$\Delta^2 \mathbf{v} = 0 \tag{14.26}$$

can be computed. Specifying boundary conditions for isolated points of deformable surfaces in principle leads to the theory of weak solutions and the associated mathematical framework for the minimization problem. The solution of the set of Eq. (14.26) belongs to the Sobolev space and is, therefore, a weak solution. It is a smooth surface that is as close as possible to a sphere and interpolates the given points.

Given a total number of M user-supplied anchor points $P_i(4 \leq i \leq M$, non-coplanar) and the normal vector at their locations, the system of equations reduces to

$$K^* \dot{\vec{v}}^* = F^*_{\vec{v}^*}, \tag{14.27}$$

where $\vec{v}^*$ stands for either $\vec{v}^{(1)^*}$, $\vec{v}^{(2)^*}$, or $\vec{v}^{(3)^*}$, the reduced vectors of the three coordinate functions, and K^* for an $(N - M) \times (N - M)$ sparse matrix that is now invertible and can be solved using a sparse linear solver. Closed 3-D objects can be initialized by selecting at least four non-coplanar points. Of course, since $F^*_{\vec{v}^*}$ depends on the surface's current position, Eq. (14.27) cannot, in general, be solved in closed form.

The algorithm for the approximation of the underlying image data is analogous to the 2-D case. Starting from the initial shape that is approximately correct in the neighborhood of the selected anchor points, the image potential is taken into account progressively for all surface vertices.

14.5.2 Model-Based Initialization

The previously described model-based approaches employing statistical encoding of large organ populations can also be successfully applied to efficient initialization of interactive methods [61]. The underlying idea is to apply statistical shape analysis for examining the remaining variability of shape due to interactive point-wise subtraction of variation. The key element is the optimal selection of principal landmarks that carry as much shape information as possible. The goal is to remove as much variation as possible by selecting points that have a maximal reduction potential. The overall process will be described below, considering the previously mentioned population of 71 hand segmented corpus callosi.

Similar to the automatic approach, the first step is the generation of a compact statistical shape description of all object instances in the database. First, we calculate the mean shape $\overline{\mathbf{p}}$ and the instance specific difference vector $\Delta\mathbf{p}_i = \mathbf{p}_i - \overline{\mathbf{p}}$.

To find the eigensystem of our data, the difference vectors are projected into a lower dimensional space whose basis M is constructed by the Gram–Schmidt

orthonormalization χ:

$$M = [\mathbf{m}_1, \ldots, \mathbf{m}_{N-1}] = \chi\,(\Delta\mathbf{p}_1, \ldots, \Delta\mathbf{p}_{N-1}) \qquad \Delta\tilde{\mathbf{p}}_i = M^T \Delta\mathbf{p}_i \quad (14.28)$$

The covariance matrix $\tilde{\Sigma}$ and the resulting PCA given by the eigensystem of $\tilde{\Sigma}$ can subsequently be calculated according to:

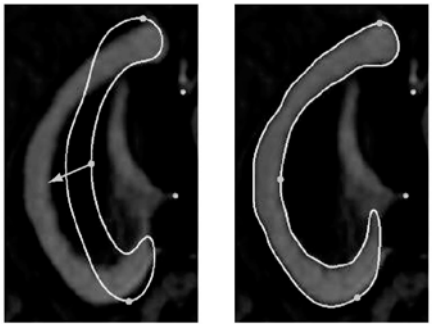
$$\tilde{\Sigma} = \frac{1}{N-1} \sum_{i=1}^{N} \Delta\tilde{\mathbf{p}}_i \Delta\tilde{\mathbf{p}}_i^T \overset{\text{PCA}}{=} \tilde{U} \Lambda \tilde{U}^T \quad \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{N-1}) \quad (14.29)$$

The principal components defining the eigenmodes in shape space are then given by back projecting the eigenvectors $\tilde{U}$:

$$U = [\mathbf{u}_1, \ldots, \mathbf{u}_{N-1}] = M\tilde{U} \quad (14.30)$$

14.5.2.1 Point-wise Subtraction of Variation

After the statistical analysis of the anatomical shape, this information can be used to progressively eliminate variation by point-wise fixation of control points. After defining the coordinate system with the AC–PC line, the initialization starts with the average model $\overline{\mathbf{p}}$ (Fig. 14.13(a)). Additional boundary conditions are then introduced by moving control vertices to approximately correct positions



(a) Initial average model and correct segmentation

(b) Basis vectors R_j

Figure 14.13: (a) Boundary conditions for an initial outline are established by prescribing a position for each coarse control vertex. (b) Shape variations caused by adding the basis vectors defining the x- and y-translation of one point to the average model. The various shapes are obtained by evaluating $\overline{\mathbf{p}} + \omega U \mathbf{r}_k$ with $\omega \in \{-2, \ldots, 2\}$ and $k \in \{x_j, y_j\}$.

on the object border. In the next step, given the *a priori* shape knowledge and these constraints, the most natural initialization outline should be chosen. In the context of PCA, this means choosing the model with minimal Mahalanobis distance D_m.

The solution to this task is to find two vectors in variation space describing decoupled x- and y-translations of a given point j in object space with minimal overall variations. Once these vectors are found, all possible boundary conditions can be satisfied by adding these appropriately weighted vectors to the mean shape.

Let $\mathbf{r}_{x_j}$ and $\mathbf{r}_{y_j}$ denote the two unknown basis vectors causing unit x- and y-translation of the point j respectively. The D_m of these two vectors is then given by

$$D_m(\mathbf{r}_k) = (\tilde{U}\mathbf{r}_k)^T \tilde{\Sigma}^{-1} \tilde{U}\mathbf{r}_k = \mathbf{r}_k^T \Lambda^{-1} \mathbf{r}_k = \sum_{e=1}^{N-1} \frac{\left(r_k^{[e]}\right)^2}{\lambda_e} \quad k \in \{x_j, y_j\} \tag{14.31}$$

Taking into account that x_j and y_j depend only on two rows of U, we define the submatrix U_j according to the following expression:

$$\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} \overline{x}_j \\ \overline{y}_j \end{bmatrix} + \begin{bmatrix} u_{2j-1\,\circ} \\ u_{2j\,\circ} \end{bmatrix} \mathbf{b} = \begin{bmatrix} \overline{x}_j \\ \overline{y}_j \end{bmatrix} + U_j \mathbf{b} \quad u_{j\,\circ} := \text{jth row of } U \tag{14.32}$$

In order to minimize D_m subject to the constraint of a separate x- or y-translation by one unit, we establish the Lagrange function L:

$$\begin{aligned} L(\mathbf{r}_k, \mathbf{l}_k) &= \sum_{e=1}^{N-1} \frac{\left(r_k^{[e]}\right)^2}{\lambda_e} - \mathbf{l}_k^T [U_j \mathbf{r}_k - \mathbf{e}_k], \\ k \in \{x_j, y_j\} \qquad \mathbf{e}_{x_j} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{e}_{y_j} = \begin{matrix} 0 \\ 1 \end{matrix} \end{aligned} \tag{14.33}$$

The vectors $\mathbf{l}_{x_j}$ and $\mathbf{l}_{y_j}$ denote the required Lagrange multipliers. To find the optimum of $L(\mathbf{r}_k, \mathbf{l}_k)$, we calculate the derivatives with respect to all elements of $\mathbf{r}_{x_j}$, $\mathbf{r}_{y_j}$, $\mathbf{l}_{x_j}$, and $\mathbf{l}_{y_j}$ and set them equal to zero:

$$\begin{aligned} \frac{\delta}{\delta \mathbf{r}_{x_j}} L(\mathbf{r}_{x_j}, \mathbf{l}_{x_j}) \stackrel{!}{=} 0 \qquad \wedge \qquad \frac{\delta}{\delta \mathbf{l}_{x_j}} L(\mathbf{r}_{x_j}, \mathbf{l}_{x_j}) \stackrel{!}{=} 0, \\ \frac{\delta}{\delta \mathbf{r}_{y_j}} L(\mathbf{r}_{y_j}, \mathbf{l}_{y_j}) \stackrel{!}{=} 0 \qquad \wedge \qquad \frac{\delta}{\delta \mathbf{l}_{y_j}} L(\mathbf{r}_{y_j}, \mathbf{l}_{y_j}) \stackrel{!}{=} 0 \end{aligned} \tag{14.34}$$

$$\begin{bmatrix} \frac{2}{\lambda_1} & & \vdots & \\ & \ddots & \vdots & -U_j^T \\ & \frac{2}{\lambda_{N-1}} & \vdots & \\ \cdots & \cdots & \cdots & \cdots \\ U_j & & \vdots & 0 \end{bmatrix} \begin{bmatrix} & \vdots & \\ \mathbf{r}_{x_j} & \vdots & \mathbf{r}_{y_j} \\ & \vdots & \\ \cdots & \cdots & \cdots \\ \mathbf{l}_{x_j} & \vdots & \mathbf{l}_{y_j} \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \mathbf{0} & \vdots & \mathbf{0} \\ & \vdots & \\ \cdots & \cdots & \cdots \\ \mathbf{e}_{x_j} & \vdots & \mathbf{e}_{y_j} \end{bmatrix} \tag{14.34'}$$

If the basis vectors and the Lagrange multipliers are combined according to $R_j = [\mathbf{r}_{x_j}\, \mathbf{r}_{y_j}]$ and $L_j = [\mathbf{l}_{x_j}\, \mathbf{l}_{y_j}]$, Eq. (14.34′) can be rewritten as two linear matrix equations:

$$2\Lambda^{-1} R_j = U_j^T L_j \tag{14.35}$$

$$U_j R_j = I \tag{14.36}$$

The two basis vectors $\mathbf{r}_{x_j}$ and $\mathbf{r}_{y_j}$ (resulting from simple algebraic operations on Eqs. (14.35) and (14.36)) are then given by

$$R_j = [\mathbf{r}_{x_j}\ \mathbf{r}_{y_j}] = \Lambda U_j^T \left[U_j \Lambda U_j^T\right]^{-1} \tag{14.37}$$

While $\mathbf{r}_{x_j}$ describes the translation of x_j by one unit with constant y_j and minimal shape variation, $\mathbf{r}_{y_j}$ alters y_j correspondingly. The resulting effect caused by adding these shape-based basis vectors to the average model is illustrated in Fig. 14.13(b). The most probable shape $\check{\mathbf{p}}$ given the displacement $[\Delta x_j, \Delta y_j]^T$ for the control vertex j is consequently determined by

$$\check{\mathbf{p}} = \overline{\mathbf{p}} + U R_j \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix}. \tag{14.38}$$

After obtaining the most probable shape for a given control vertex, we now have to ensure that subsequent modifications do not alter the adjusted vertex. Therefore, we remove the components from the statistic that cause a displacement of the point. The first step is to subtract the basis vectors R_j, weighted by the example specific displacement $U_j = [\Delta x_j, \Delta y_j]_i^T$, from the parameter representation $\mathbf{b}_i$ of each instance i:

$$\mathbf{b}_i^{\hat{j}} = \mathbf{b}_i - R_j U_j \mathbf{b}_i = \left(I - R_j U_j\right) \mathbf{b}_i \quad \forall i \in \{1, \ldots, N\} \tag{14.39}$$

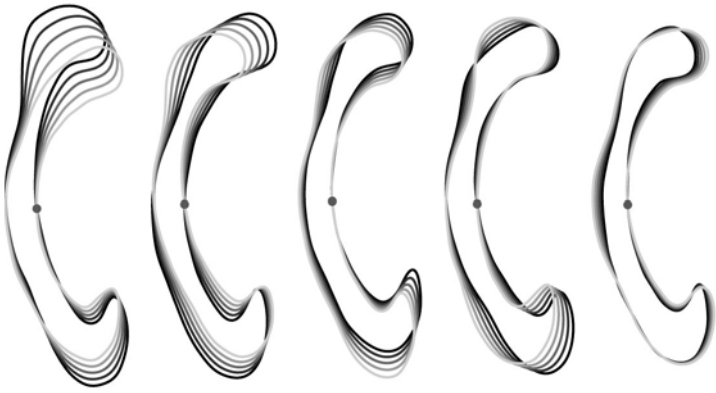


Figure 14.14: The first five one-point invariant eigenmodes after the subtraction of the first principal landmark. The various shapes are obtained by evaluating $\overline{\mathbf{p}} + \omega\sqrt{\lambda_k^{\hat{j}}}\mathbf{u}_k^{\hat{j}}$ with $\omega \in \{-2, \ldots, 2\}$ and $k \in \{1, \ldots, 5\}$.

Doing so for all instances, we obtain a new description of our population $\mathbf{b}_i^{\hat{j}}$ which is invariant with respect to point j. An example of the removal of the variation is visualized in Fig. 14.14.

In order to further improve the point-wise elimination process, the control point selection strategy has to be optimized. This can be done by choosing control vertices, or *principal landmarks*, which carry as much shape information as possible.

We define the reduction potential of a vertex j_k being a candidate to serve as the kth principal landmark by

$$P(j_k) = -\sum_{l=1}^{N-1-2(k-1)} \left(\tilde{\sigma}_l^2\right)^{\hat{s}_k} = -\mathrm{tr}(\tilde{\Sigma}^{\hat{s}_k}) = -\mathrm{tr}(\Lambda^{\hat{s}_k}), \tag{14.40}$$

with sequence $\hat{s}_k = \{\hat{j}_1, \ldots, \hat{j}_k\}$ denoting the set of the k point-indices of the principal landmarks that have been removed from the statistic in the given order, and the superscript $\circ^{\hat{s}_k}$ indicating the value of $\circ$ if the principal landmarks $\hat{s}_k$ have been removed.

In order to remove as much variation as possible, we choose consequently that point as the first principal landmark that holds the largest reduction potential: $j_1 = \max_j[P(j)]$. This selection strategy was applied to obtain the eigenmodes shown in Fig. 14.14. Further application of the selection strategy

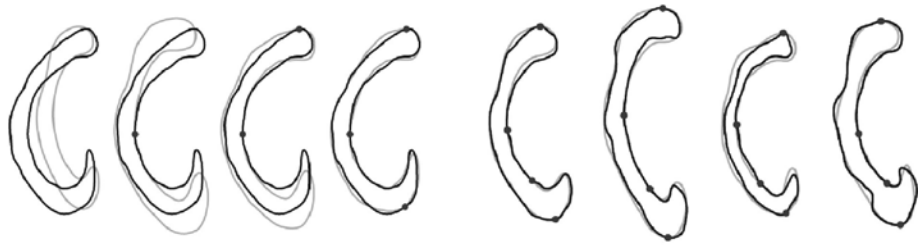
(a) Two-point invariant eigenmodes (b) Three-point invariant

Figure 14.15: Remaining variability after vertex elimination of (a) two and (b) three principal landmarks.

to the example, obtains the optimal second and third principal landmark (Fig. 14.15).

14.5.2.2 Initialization Process

The described framework can now be used for efficient initialization of deformable models. Examples of the initialization process are shown in Fig. 14.16. The left image shows how the initial average model converges toward a sound approximation by adding control vertices. The right image depicts four additional examples with adjusted principle landmarks. Generally speaking,



(a) Progressive initialization (b) Four different initializations

Figure 14.16: (a) Generation of an initial outline for segmentation. Shape instance in black and fitted initializations in gray with an increasing number of fitted principal landmarks. (b) Initial shapes with four adjusted principal landmarks for the segmentation of four randomly chosen instances.

selecting three to four landmarks has proven to be sufficient for a reasonably good initialization.

14.6 Improving Human–Computer Interaction

14.6.1 Background

Extensive research has been invested in recent years into improving interactive segmentation algorithms. It is, however, striking that the human–computer interface, a substantial part of an interactive setup, is usually not investigated. Although the need for understanding the influence of human–computer interaction on interactive segmentation is recognized, only very little research has been done in this direction.

In order to improve information flow and to achieve optimal cooperation between interactive image analysis algorithms and human operators, we evaluated closed-loop systems utilizing new man–machine interfacing paradigms [62].

The mouse-based, manual initialization of deformable models in two dimensions represents a major bottleneck in interactive segmentation. In order to overcome the limitations of 2-D viewing and interaction the usage of direct 3-D interaction is inevitable. However, adding another dimension to user interaction causes several problems. Editing, controlling, and interacting in three dimensions often overwhelms the perceptual powers of a human operator. Furthermore, today's desktop metaphors are based on 2-D interaction and cannot easily be extended to the volumetric case. Finally, the visual channel of the human sensory system is not suitable for the perception of volumetric data.

However, these major drawbacks are valid only in terms of interactive systems that are based on 2-D Window–Icon–Mouse–Pointer (WIMP) interfaces that solely rely on the visual sense of the human operator. In order to alleviate the limitations of visual-only systems we may try to enhance the interaction process with additional sensory feedback. The fundamental challenge here is to find efficient ways for information flow between user and computer. Several sensory channels could be addressed, but due to the 3-D nature of the problem, the most obvious choice is the haptic channel. As an example, a multimodal system using visual and haptic volumetric rendering will be described, which was successfully applied to the segmentation of the intestinal system (Fig. 14.17).

Figure 14.17: Interactive, multimodal setup.

14.6.2 Multimodal Segmentation

The validity of the multimodal approach is evaluated on the basis of the highly complex task of segmentation of the small intestine. Currently, no satisfying solutions exist to deal with this problem. Although the small intestine has a complex spatial structure, from a topological point of view it is a rather simple linear tube with exactly defined start- and endpoint. Therefore, the key to solving the segmentation problem is to make use of the topological causality of the structure. The overall extraction process has to be mapped onto the linear structure of the intestinal system to simplify the task.

The initial step of our multimodal technique is the haptically assisted extraction of the centerline of the tubular structures. The underlying idea is to create guiding force maps, similar to the notion of virtual fixtures found in teleoperation [63, 64]. These forces can be used to assist a user's movement through the complex dataset.

To do this we first create a binarization of our data volume V by thresholding. The threshold is chosen dependent on the grayscale histogram, but can also be specified manually. We have to emphasize that this step is not sufficient for a complete segmentation of the datasets we are interested in. This is due to the often low quality of the image data. Nevertheless, in the initial step we are not interested in a topologically correct extraction. On the contrary, we only need a rough approximation of our object of interest. From the resulting dataset W

we generate an Euclidean distance map by computing the value

$$DM(x, y, z) = \min_{(x_i, y_i, z_i) \in \overline{W}} d[(x, y, z), (x_i, y_i, z_i)], \tag{14.41}$$

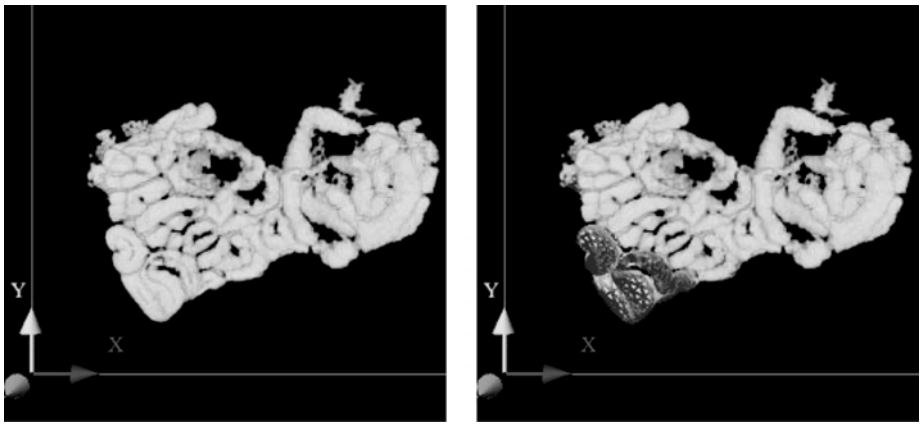
for each $(x, y, z) \in W$, where d denotes the Euclidean distance from a voxel that is part of the tubular structure to a voxel of the surrounding tissue $\overline{W} = V \setminus W$.

In the next step we negate the 3-D distance map and approximate the gradients by central differences. Moreover, to ensure the smoothness of the computed forces, we apply a $5 \times 5 \times 5$ binomial filter. This force map is precomputed before the actual interaction to ensure a stable force-update. Because the obtained forces are located at discrete voxel positions, we have to do a trilinear interpolation to obtain the continuous gradient force map needed for stable haptic interaction. Furthermore, we apply a low-pass filter in time to further suppress instabilities. The computed forces can now be utilized to guide a user on a path close to the centerline of the tubular structure. In the optimal case of good data quality, the user falls through the dataset guided along the 3-D ridge created by the forces. However, if the 3-D ridge does not exactly follow the centerline the user can guide the 3-D cursor by exerting a gentle force on the haptic device to leave the precalculated curve.

While moving along the path, points near the centerline are set. These points can be used to obtain a B-spline, which approximates the path. In the next step this extracted centerline is used to generate a good initialization for a deformable surface model. To do this, a tube with varying thickness is created according to the precomputed distance map. This resulting object is then deformed subject to a thin plate under tension model. Details of the algorithmic background of this deformable model approach are described in section 14.4.2.

Because of the good initialization, only a few steps are needed to approximate the desired object [65]. The path initialization can be seen in Fig. 14.18(a). Note, that the 3-D data is rendered semitransparent to visualize the path in the lower left portion of the data. Figure 14.18(b) depicts the surface model during deformation.

In order to further improve the interaction with complicated datasets a step-by-step segmentation approach can be adopted by hiding already segmented loops. This allows a user to focus attention on the parts that still have to be extracted. For this purpose the 3-D surface model is turned back into voxels and removed from the dataset (Fig. 14.19). This step can be carried out in

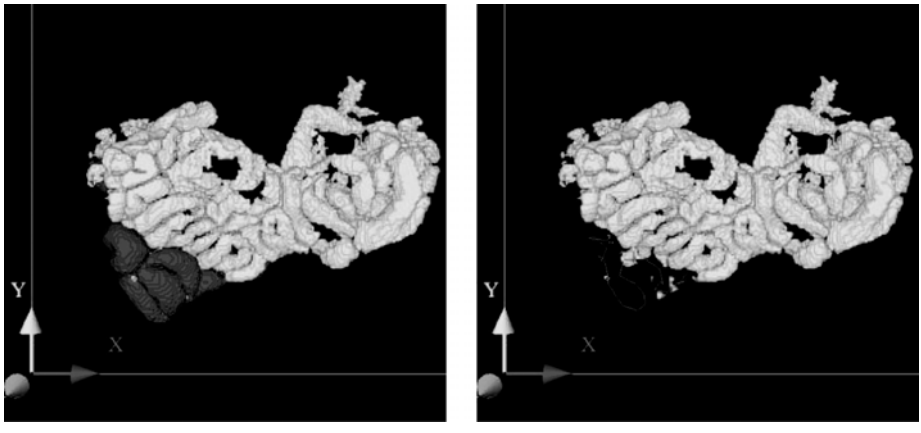


(a) Intialized path. (b) Deforming tube.

Figure 14.18: Interactive segmentation.

real-time by using a hardware accelerated, z-buffer based approach as described in [66].

In order to validate the described system, it was used to generate topologically correct models of the small intestine. The application studies of the system required interaction times of 20–30 min, which compares favorably to the reported times of 1–2 h in previous research. It was possible to extract the centerlines of the complicated datasets, obtain the segmentations, and create



(a) Voxelization. (b) Removed segemented part.

Figure 14.19: Hiding segmented parts.

virtual fly-throughs. Further evaluation studies were performed, which show a statistically significant performance improvement in the trial time when using haptically enhanced interaction in 3-D segmentation. Also in the haptic condition the quality of segmentation was always superior to the one without force feedback.

14.7 Conclusions

In spite of the enormous research and development effort invested into finding satisfactory solutions during the past decades, the problem of medical image segmentation (as image segmentation in general) is still an unsolved problem today, and no single approach is able to successfully address the whole range of possible clinical problems. The basic reason for this rather disappointing status lies in the difficulties in representing and using the prior information in its full extent, which is necessary to successfully solve the underlying task in scene analysis and image interpretation.

While first results already clearly demonstrate the power of the model-based techniques, generic segmentation systems capable to analyze a broad range of radiological data even under severely pathological conditions cannot be expected in the near future. Currently available methods, like those discussed in this chapter, allow to work only within a very narrow, specialized problem domain and fundamental difficulties have to be expected if trying to establish more generic platforms. The practically justifiable number of examples in the training sets can cover only very limited variations of the anatomy and are usually applied to analyzing images without large pathological changes. It still needs a long way to go, before the computer representation and usage of the prior knowledge involved in the interpretation of radiological images can be represented and used by a computer in complexity which is sufficient to reasonably imitate the everyday work of an experienced clinical radiologist. Accordingly, in the near future only a well-balanced cooperation between computerized image analysis methods and a human operator will be able to efficiently address many clinically relevant segmentation problems. Better understanding of the perceptional and technical principles of man–machine interaction is therefore a fundamentally important research area which should now get significantly more attention than what it was getting in the past.

Bibliography

[1] Gerig, G., Martin, J., Kikinis, R., Kübler, O., Shenton, M., and Jolesz, F., Automatic segmentation of dual-echo MR head data, In: Proceedings of Information Processing in Medical Imaging'91, Wye, GB, 1991, pp. 175–187.

[2] Duda, R. and Hart, P., Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[3] Shattuck, D., Sandor-Leahy, S., Schaper, K., Rottenberg, D., and Leahy, R., Magnetic resonance image tissue classification using a partial volume model, Neuroimage, Vol. 13, pp. 856–876, 2001.

[4] S. Ruan, J. X., C. Jaggi and Bloyet, J., Brain tissue classification of magnetic resonance images using partial volume modeling, IEEE Trans. Med. Imaging, Vol. 19, No. 12, pp. 172–186, 2000.

[5] Gerig, G., Kübler, O., Kikinis, R., and Jolesz, F., Nonlinear anisotropic filtering of MRI data, IEEE Trans. Med. Imaging, Vol. 11, No. 2, pp. 221–232, 1992.

[6] Guillemaud, R. and Brady, M., Estimating the bias field of MR images, IEEE Trans. Med. Imaging, Vol. 16, No. 3, pp. 238–251, 1997.

[7] M. Styner, G. S., Ch. Brechbühler and Gerig, G., Parametric estimate of intensity inhomogeneities applied to MRI, IEEE Trans. Med. Imaging, Vol. 19, No. 3, pp. 153–165, 2000.

[8] Wells, W., Grimson, W., Kikinis, R., and Jolesz, F., Adaptive segmentation of MRI data, IEEE Trans. Med. Imaging, Vol. 15, No. 4, pp. 429–443, 1996.

[9] Van Leemput, K., Maes, F., Bello, F., Vandermeulen, D., Colchester, A., and Suetens, P., Automated segmentation of MS lesions from multichannel MR images, In: Proceedings of Second International Conference on Medical Image Computing and Computer-Assisted Interventions, MICCAI'99, Taylor, C. and Colchester, A., eds., Lecture Notes in Computer Science, Vol. 1679, Springer-Verlag, New-York, pp. 11–21, 1999.

[10] Li, S. Z., Markov Random Field Modeling in Computer Vision, Springer-Verlag, Tokyo, 1995.

[11] Serra, J., Image Analysis and Mathematical Morphology, Academic Press, San Diego, 1982.

[12] Raya, S., Low-level segmentation of 3-D magnetic resonance brain images—A rule-based system, IEEE Trans. Med. Imaging, Vol. 9, No. 3, pp. 327–337, 1990.

[13] Stansfield, S., ANGY: A rule-based system for automatic segmentation of coronary vessels from digital subtracted angiograms, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 8, No. 2, pp. 188–199, 1986.

[14] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching, Comput. Vision Graph. Image Process., Vol. 46, pp. 1–21, 1989.

[15] Evans, A. C., Collins, D. L., and Holmes, C. J., Toward a probabilistic atlas of human neuroanatomy, In: Brain Mapping: The Methods, Mazziotta, J. C. and Toga, A. W., eds., Academic Press ISBN 0126930198 pp. 343–361, 1996.

[16] Jiang, H., Holton, K., and Robb, R., Image registration of multimodality 3-D medical images by chamfer matching, In: Proceedings of Biomedical. Image Processing and 3D Microscopy, SPIE, Vol. 1660, SPIE, The International Society of Optical Engineering pp. 356–366, 1992.

[17] Christensen, G., Miller, M., and Vannier, M., Individualizing neuroanatomical atlases using a massively parallel computer, IEEE Computer, pp. 32–38, January 1996.

[18] Bookstein, F., Shape and the information in medical images: A decade of the morphometric synthesis, Comput. Vision. Image Understand., Vol. 66, No. 2, pp. 97–118, 1997.

[19] Evans, A., Kamber, M., Collins, D., and MacDonald, D., An MRI-based probabilistic atlas of neuroanatomy, In: Magnetic Resonance Scanning and Epilepsy, Shorvon, S., ed., Plenum Press, New York, pp. 263–274, 1994.

[20] Wang, Y. and Staib, L., Elastic model based non-rigid registration incorporating statistical shape information, In: Proc. First Int. Conf. on Medical Image Comp. and Comp. Assisted Interventions, Vol. 1679 of Lecture Notes in Comp. Sci., pp. 1162–1173, Springer-Verlag, New York, 1998.

[21] Terzopoulos, D. and Metaxas, D., Dynamic 3D models with local and global deformations: Deformable superquadrics, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 13, No. 7, pp. 703–714, 1991.

[22] Vemuri, B. and Radisavljevic, A., Multiresolution stochastic hybrid shape models with fractal priors, ACM Trans. Graphics, Vol. 13, No. 2, pp. 177–200, 1994.

[23] Staib, L. and Duncan, J., Boundary finding with parametrically deformable models, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, No. 11, pp. 1061–1075, 1992.

[24] Brechbühler, C., Gerig, G., and Kübler, O., Parametrization of closed surfaces for 3-D shape description, CVGIP: Image Understand., Vol. 61, pp. 154–170, 1995.

[25] Cootes, T., Cooper, D., Taylor, C., and Graham, J., Training models of shape from sets of examples, In: Proceedings of The British Machine Vision Conference (BMVC) Springer-Verlag, New-York, pp. 9–18, 1992.

[26] Cootes, T. and Taylor, C., Active shape models—'Smart snakes,' In: Proceedings of The British Machine Vision Conference (BMVC) Springer-Verlag, New-York, pp. 266–275, 1992.

[27] Rangarajan, A., Chui, H., and Bookstein, F., The softassign procrustes matching algorithm, information processing in medical imaging, pp. 29–42, 1997. Available at http://noodle.med.yale.edu/anand/ps/ipsprfnl.ps.gz.

[28] Tagare, H., Non-rigid curve correspondence for estimating heart motion, Inform. Process. Med. Imaging, Vol. 1230, pp. 489–494, 1997.

[29] Kotcheff, A. and Taylor, C., Automatic construction of eigenshape models by genetic algorithm, Inform. Process. Med. Imaging, Vol. 1230, pp. 1–14, 1997.

[30] Kelemen, A., Szekely, G., and Gerig, G., Elastic model-based segmentation of 3-d neuroradiological data sets, IEEE Trans. Med. Imaging, Vol. 18, pp. 828–839, 1999.

[31] Staib, L. and Duncan, J., Model-based deformable surface finding for medical images, IEEE Trans. Med. Imaging, Vol. 15, No. 5, pp. 1–12, 1996.

[32] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J., Active shape models—Their training and application, Comput. Vision Image Understand., Vol. 61, No. 1, pp. 38–59, 1995.

[33] Székely, G., Kelemen, A., Brechbühler, C., and Gerig, G., Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models, Med. Image Anal., Vol. 1, No. 1, pp. 19–34, 1996.

[34] McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A survey, Med. Image Anal., Vol. 1, No. 2, pp. 91–108, 1996.

[35] Cootes, T., Edwards, G., and Taylor, C., Active appearance models, In: Proceedings of the European Conference on Computer Vision, Vol. 2, Springer-Verlag, New-York, pp. 484–498, 1998.

[36] Kelemen, A., Szekely, G., and Gerig, G., Elastic model-based segmentation of 3-d neuroradiological data sets, IEEE Trans. Med. Imaging, Vol. 18, No. 10, pp. 828–839, 1999.

[37] Kruggel, F. and Lohmann, G., Automatical adaption of the stereotactical coordinate system in brain MRI datasets, In: Information Processing in Medical Imaging, Springer-Verlag, New York, pp. 471–476, 1997.

[38] Barrett, W. and Mortensen, E., Interactive live-wire boundary extraction, Medical Image Analysis, pp. 331–341, 1997. Available at citeseer.nj.nec.com/barrett97interactive.html.

[39] Fischler, M., Tenenbaum, J., and Wolf, H., Detection of roads and linear structures in low-reslution aerial imagery using a multisource knowledge integration technique, Comput. Graph. Image Process., Vol. 15, pp. 201–233, 1981.

[40] O'Donnell, L., Weslin, C.-T., Grimson, W. E. L., Ruiz-alzola, J., Shenton, M. E., and Kikinis, R., Phase-based user-steered image segmentation, In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2001, pp. 1022–1030.

[41] Falcão, A. and Udapa, J., A 3D generalization of user-steered live-wire segmentation, Med. Image Anal., Vol. 4, No. 1, pp. 389–402, 1997.

[42] Falcão, A., Udapa, J., and Miyazawa, F., An ultra-fast user-steered image segmentation paradigm: Live wire on the fly, IEEE Trans. Med. Imaging, Vol. 19, No. 1, pp. 55–62, 2000.

[43] Haenselmann, T. and Effelsberg, W., Wavelet-based semi-automatic live-wire segmentation, SPIE Human Vision and Electronic Imaging VIII, Vol. 5007, pp. 260–269, 2003. Available at citeseer.nj.nec.com/569760.html.

[44] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput. Vision, Vol. 1, No. 4, pp. 321–331, 1988.

[45] Canny, J., A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, No. 6, pp. 679–698, 1986.

[46] Fua, P. and Leclerc, Y., Model driven edge detection, Mach. Vision Appl., Vol. 3, pp. 45–56, 1990.

[47] Leymarie, F. and Levine, M., Tracking deformable objects in the plane using an active contour model, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 15, No. 6, pp. 617–634, 1993.

[48] Samadani, R., Changes in connectivity in active contour models, In: Proceedings of the IEEE Workshop on Visual Motion, Irvine, California, March 1989, pp. 337–343.

[49] Terzopoulos, D., On matching deformable models to images, Topical Meeting Mach. Vision Tech. Digest Series, Vol. 12, pp. 160–167, 1987.

[50] Cohen, L. and Cohen, I., A finite element method applied to new active contour models and 3D reconstructions, In: Proceedings of the Third International Conference on Computer Vision, Osaka, Japan, Dec. 1990, pp. 587–591.

[51] Cohen, I., Cohen, L. D., and Ayache, N., Using deformable surfaces to segment 3-D images and infer differential structures, Comput. Vision Graph. Image Process., Vol. 56, No. 2, pp. 242–263, 1992.

[52] Hug, J., Brechbühler, C., and Székely, G., Tamed snake: A particle system for robust semi-automatic segmentation, In: MICCAI, 1999, pp. 106–115.

[53] Dyn, N., Levin, D., and Gregory, J., A 4-point interpolatory subdivision scheme for curve design, Comput. Aided Geomet. Design, Vol. 4, No. 4, pp. 257–268, 1987.

[54] Hug, J., Semi-Automatic Segmentation of Medical Imagery, Ph.D. Thesis, ETH Zürich-Swiss Federal Institute of Technology, 2001.

[55] Dyn, N., Levin, D., and Gregory, J., A butterfly subdivision scheme for surface interpolation with tension control, Trans. Graph., Vol. 9, No. 2, pp. 160–169, 1990.

[56] Zorin, D., Schröder, P., and Sweldens, W., Interpolating subdivision for meshes of arbitrary topology, In: SIGGRAPH, August 1996, pp. 189–192.

[57] Kobbelt, L., Iterative Erzeugung glatter Interpolatoren., Ph.D. Thesis, University at Karlsruhe, 1994.

[58] Schneider, R. and Kobbelt, L., Geometric fairing of irregular meshes for free-form surface design, Comput. Aided Geomet. Design, Vol. 18, No. 4, pp. 359–379, 5 2001.

[59] Desbrun, M., Meyer, M., Schroder, P., and Barr, A., Discrete Differential-Geometry Operators in nD, preprint, The Caltech Multi-Res Modeling Group, 2000.

[60] Neuenschwander, W., Fua, P., Székely, G., and Kübler, O., Initializing snakes, In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 1994, pp. 658–663.

[61] Hug, J., Brechbühler, C., and Székely, G., Model-based initialisation for segmentation, In: Proceedings of 6th European Conference on Computer Vision (ECCV 2000), Part II, Vernon, D., ed., Lecture Notes in Computer Science, Springer, Berlin pp. 290–306, 2000.

[62] Harders, M. and Székely, G., Enhancing human computer interaction in medical segmentation, Proc. IEEE, Vol. 91, No. 9, pp. 1430–1442, 2003.

[63] Rosenberg, L., Virtual fixtures: Perceptual tools for telerobotic manipulation, In: IEEE Virtual Reality Annual International Symposium, 1993, pp. 76–82.

[64] Sayers, C. and Paul, R., An operator interface for teleprogramming employing synthetic fixtures, Presence Teleoperat. Virtual Environ., Vol. 3, pp. 309–320, 1994.

[65] Harders, M. and Székely, G., New paradigms for interactive 3D volume segmentation, J. Visual. Comput. Animation, Vol. 13, pp. 85–95, 2002.

[66] Karabassi, E.-A., Papaioannou, G., and Theoharis, T., A fast depth-buffer-based voxelization algorithm, J. Graph. Tools, Vol. 4, No. 4, pp. 5–10, 1999.

The Editors

Dr. Jasjit S. Suri received his BS in computer engineering with distinction from Maulana Azad College of Technology, Bhopal, India, his MS in computer sciences from University of Illinois, Chicago, and Ph.D. in electrical engineering from University of Washington, Seattle. He has been working in the field of computer engineering/imaging sciences for 20 years. He has published more than 125 technical papers in body imaging. He is a lifetime member of research engineering societies: Tau-Beta Pi, Eta-Kappa-Nu, Sigma-Xi, and a member of NY Academy of Sciences, Engineering in Medicine and Biology Society (EMBS), SPIE, ACM, and is also a senior member at IEEE. He is in the editorial board/reviewer of several international journals such as *Real Time Imaging*, *Pattern Analysis and Applications*, *Engineering in Medicine and Biology*, *Radiology*, *Journal of Computer Assisted Tomography*, *IEEE Transactions of Information Technology in Biomedicine*, *and IASTED Board.*

He has chaired image processing tracks at several international conferences and has given more than 40 international presentations/seminars. Dr. Suri has written four books in the area of body imaging (such as cardiology, neurology, pathology, mammography, angiography, atherosclerosis imaging) covering medical image segmentation, image and volume registration, and physics of medical imaging modalities like: MRI, CT, X-ray, PET, and ultrasound. He also holds several United States patents. Dr. Suri has been listed in Who's Who seven times, is a recipient of president's gold medal in 1980, and has received more than 50 scholarly and extracurricular awards during his career. He is also a Fellow of American Institute of Medical and Biological Engineering (AIMBE) and ABI. Dr. Suri's major interests are: computer vision, graphics and image processing (CVGIP), object oriented programming, image guided surgery and teleimaging. Dr. Suri had worked for Philips Medical Systems and Siemens Medical Research Divisions. He is also a visiting professor with the department of computer science, University of Exeter, Exeter, England. Currently, Dr. Suri is with JWT Inc. as director of biomedical engineering division (in opthalmology imaging) in conjunction with Biomedical Imaging Laboratories, Case Western Reserve University, Cleveland.

Dr. David Wilson is a professor of biomedical engineering and radiology, Case Western Reserve University. He has research interests in image analysis, quantitative image quality, and molecular imaging, and he has a significant track record of federal research funding in these areas. He has over 60 refereed journal publications and has served as a reviewer for several leading journals. Professor Wilson has six patents and two pending patents in medical imaging. Professor Wilson has been active in the development of international conferences; he was Track Chair at the 2002 EMBS/BMES conference, and he was Technical Program Co-Chair for the 2004 IEEE International Symposium on Biomedical Imaging.

Professor Wilson teaches courses in biomedical imaging, and biomedical image processing and analysis. He has advised many graduate and undergraduate students, all of whom are quite exceptional, and has been primary research advisor for over 16 graduate students since starting his academic career. Prior to joining CWRU, he worked in X-ray imaging at Siemens Medical Systems at sites in New Jersey and Germany. He obtained his PhD from Rice University. Professor Wilson has actively developed biomedical imaging at CWRU. He has led a faculty recruitment effort, and he has served as PI or has been an active leader on multiple research and equipment developmental awards to CWRU, including an NIH planning grant award for an In Vivo Cellular and Molecular Imaging Center and an Ohio Wright Center of Innovation award. He can be reached at dlw@po.cwru.edu.

Dr. Swamy Laxminarayan currently serves as the chief of biomedical information engineering at the Idaho State University. Previous to this, he held several senior positions both in industry and academia. These have included serving as the chief information officer at the National Louis University, director of the pharmaceutical and health care information services at NextGen Internet (the premier Internet organization that spun off from the NSF sponsored John von Neuman National Supercomputer Center in Princeton), program director of biomedical engineering and research computing and program director of computational biology at the University of Medicine and Dentistry in New Jersey, vice-chair of Advanced Medical Imaging Center, director of clinical computing at the Montefiore Hospital and Medical Center and the Albert Einstein College of Medicine in New York, director of the VocalTec High Tech Corporate University

in New Jersey, and the director of the Bay Networks Authorized Center in Princeton. He has also served as an adjunct professor of biomedical engineering at the New Jersey Institute of Technology, a clinical associate professor of health informatics, visiting professor at the University of Bruno in Czech Republic, and an honorary professor of health sciences at Tsinghua University in China.

As an educator, researcher, and technologist, Prof. Laxminarayan has been involved in biomedical engineering and information technology applications in medicine and health care for over 25 years and has published over 250 scientific and technical articles in international journals, books, and conferences. His expertize lies in the areas of biomedical information technology, high performance computing, digital signals and image processing, bioinformatics, and physiological systems analysis. He is the coauthor of the book *State-of-the-Art PDE and Level Sets Algorithmic Approaches to Static and Motion Imagery Segmentation* published by Kluwer Publications and the book *Angiography Imaging: State-of-the-Art-Acquisition*, *Image Processing and Applications Using Magnetic Resonance*, *Computer Tomography*, *Ultrasound and X-ray*, *Emerging Mobile E-Health Systems* published by the CRC Press and two volumes of *Handbook of Biomedical Imaging* to be published by Kluwer Publications. He has also worked as the editor/coeditor of 20 international conferences and has served as a keynote speaker in international conferences in 13 countries.

He is the founding editor-in-chief and editor emeritus of IEEE Transactions on Information Technology in Biomedicine. He served as an elected member of the administrative and executive committees in the IEEE Engineering in Medicine and Biology Society and as the society's vice president for 2 years. His other IEEE roles include his appointments as program chair and general conference chair of about 20 EMBS and other IEEE conferences, an elected member of the IEEE Publications and Products Board, member of the IEEE Strategic Planning and Transnational Committees, member of the IEEE Distinguished Lecture Series, delegate to the IEEE USA Committee on Communications and Information Policy (CCIP), U.S. delegate to the European Society for Engineering in Medicine, U.S. delegate to the General Assembly of the IFMBE, IEEE delegate to the Public Policy Commission and the Council of Societies of the AIMBE, fellow of the AIMBE, senior member of IEEE, life member of Romanian Society of Clinical Engineering and Computing, life member of Biomedical Engineering Society of India, and U.S. delegate to IFAC and IMEKO Councils in TC13. He was recently elected to the Administrative Board of the International Federation for

Medical and Biological Engineering, a worldwide organization comprising 48 national members, overseeing global biomedical engineering activities. He was also elected to serve as the publications co-chairman of the Federation.

His contributions to the discipline have earned him numerous national and international awards. He is a fellow of the American Institute of Medical and Biological Engineering, a recipient of the IEEE 3rd Millennium Medal, and a recipient of the Purkynje award from the Czech Academy of Medical Societies, a recipient of the Career Achievement Award, numerous outstanding accomplishment awards, and twice recipient of the IEEE EMBS distinguished service award. He can be reached at s.n.laxminarayan@ieee.org.

Index