Statistical Shape Characterization Using the Medial Representation

by

Paul A. Yushkevich

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2003

Approved by

Advisor: Professor Stephen M. Pizer, Ph.D.

Reader: Professor Stephen R. Aylward, Ph.D.

Reader: Professor Guido Gerig, Ph.D.

Reader: Professor Sarang Joshi, Ph.D.

Reader: Professor J. S. Marron, Ph.D.

ABSTRACT

PAUL A. YUSHKEVICH: Statistical Shape Characterization
Using the Medial Representation
(Under the direction of Stephen M. Pizer, Ph.D.)

The goal of the research presented in this dissertation is to improve the clinical understanding of processes that affect the shape of anatomical structures. Schizophrenia is an example of such a process: it is known to affect the shape of the hippocampus, but the precise nature of the morphological changes that it causes is not fully understood. This dissertation introduces novel statistical shape characterization methodology that can improve the understanding of shape-altering biological processes by (i) identifying the regions of the affected objects where these processes are most significantly manifested and (ii) expressing the effects of these processes in intuitive geometric terms. The following three new techniques are described and evaluated in this dissertation.

1. In an approach motivated by human form perception, the shape characterization problem is divided into a coarse-to-fine hierarchy of sub-problems that analyze shape at different locations and levels of detail, making it possible to compare the effects of shape-altering processes on different object regions. Statistical features are based on the medial (skeletal) object representation, which can be used to decompose objects into simple components called figures and to measure the bending and widening of the figures. Such features make it possible to express shape variability in terms of bending and widening.

2. A new algorithm that identifies regions of biological objects that are most relevant for shape-based classification is developed. In the schizophrenia application, the algorithm is used to find the hippocampus locations most relevant for classification between schizophrenia patients and matched healthy controls. The algorithm fuses shape heuristics with existing feature selection methodology, effectively reducing the inherently combinatorial search space of the latter.

3. Biological objects in 3D and 2D are described using a novel medial representation that models medial loci and boundaries using continuous manifolds. The continuous medial representation is used in the deformable templates framework to segment objects in medical images. The representation allows arbitrary sampling that is needed by the hierarchical shape characterization method.

ACKNOWLEDGEMENT

This dissertation and the research that it describes were made possible thanks to the enthusiastic support of many people.

First and foremost, my sincere gratitude goes to my advisor Dr. Stephen M. Pizer. I thank him for the depth and breadth of his guidance, for his constant willingness to help, for his kind and supportive words at the uncertain times, and for the example that he sets in his commitment to integrity.

To only thank Dr. Stephen Aylward, Dr. Guido Gerig, Dr. Sarang Joshi, and Dr. J.S. Marron for their roles as members of my dissertation committee is to not give them enough credit. They have helped me enormously along the way by involving themselves in my research, collaborating on research projects, and co-authoring publications.

My research was shaped through interaction with many professors, fellow students and members of the academic staff. I am deeply grateful for the contributions of Dr. Edward L. Chaney, Dr. James Damon, Dr. Daniel S. Fristch, Dr. Keith E. Muller, Dr. Martin Styner, P. Thomas Fletcher, Yonatan Fridman, Sean Ho, and Andrew Thall, who are or were working at University of North Carolina, and Dr. Lei Wang and Dr. John Csernasky at the Washington University Medical School. I appreciate the essential feedback received on numerous occasions from the participants of the Image Lunch, Medial Geometry, and Shape Statistics seminars.

I extend my gratitude to the faculty, students, and staff of the Department of Computer Science at the University of North Carolina at Chapel Hill, who over the years have helped me in many ways. I am especially thankful for the organizational support and assistance of Delphine Bull.

I am indebted to the people and institutions whose support has made my research possible. My work has been funded by the Medical Image Display and Analysis Group (MIDAG) headed by Dr. Stephen M. Pizer, the Departments of Computer Science and Radiation Oncology at UNC, the National Institutes of Health and the National Library of Medicine, and the Intel Corporation. I thank the people of the Unites States of America for welcoming my family and I into this great country and for bearing much of the financial burden for my education. I bow my head to all those who have made and continue to make sacrifices in order to make it possible for people like me to pursue their dreams.

I warmly thank my friends and my family for their kindness, understanding, and support. I would not be here without my parents, and I would not be doing a Ph.D. without their relentless patience. And my world would not be complete without Reyes, the woman with whom I wish to share the rest of my life.

CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

1.1. Motivation

Recent advances in physics, biomedical engineering and computer science have led to the development of several non-intrusive, safe, and relatively affordable medical imaging techniques. These techniques produce high-resolution images that allow doctors to look inside of a living human body. Careful analysis of medical images extends their utility beyond simple visual inspection, providing answers to critical questions about human anatomy, physiology, and disease.

Among the research areas that compose the field of medical image analysis lies statistical shape characterization, which studies the variability in the geometric form of objects found in medical images. It is in this area where the contributions reported in this dissertation fall. The remainder of this section describes the motivation for three different aspects of the dissertation. While its main focus is on *anatomical shape characterization*, it also makes significant contributions to the areas of *medial object representation* and *feature selection*. The following sub-sections briefly describe these areas and the driving problems within them.

1.1.1. Anatomical Shape Characterization. The term *shape characterization* can be defined as the application of statistical methodology in order to measure and describe the variability in the geometric form of objects, which may fall in a number of known categories. Anatomical shape characterization studies biological objects, such as organs,

and computes their geometric form by analyzing medical images. This section describes the goals and driving questions of this important area of scientific research.

One motivation for anatomical shape characterization lies in its potential ability to serve as a diagnostic tool. There exist a number of diseases that are known to consistently alter the shape of certain organs. For example, it has been reported that the shape of the hippocampus undergoes changes with the progress of schizophrenia [Csernansky et al. 1998, Shenton et al. 2002]. By characterizing the shape differences between healthy and diseased hippocampi, while taking into account the normal variability in the shape of the hippocampus in the human population, it may be possible to derive a set of decision rules that can classify new instances of the hippocampus as healthy or diseased. Such rules could then be used to aid early diagnosis of schizophrenia.

Beyond providing diagnosis, shape characterization can also offer important clues about the nature of certain diseases by determining whether, where, and how the shape of different organs is affected. Although the nature of schizophrenia is not fully understood, recent evidence suggests that schizophrenia affects the shape of some parts of the hippocampus more than others. If confirmed, such evidence would allow neuroscientists to focus their research on the cells contained in that particular area of the hippocampus.

Shape characterization can also aid the medical field by providing statistical anatomical atlases. Presently, anatomical books show one typical instance of human anatomy, as well as some instances of anatomical abnormalities [Netter 1997, Talairach and Tournoux 1988, Hohne et al. 1992]. Researchers are using shape characterization techniques to estimate the variability in the shape of different parts of the human body in order to construct atlases that can show many realistic variations of human anatomy. Not only do these atlases serve educational purposes, but they also improve automatic image segmentation algorithms by imposing a prior on the shape of the objects being segmented.

To conduct a shape characterization study, researchers select a group of qualifying individuals and acquire medical images of the anatomical area of interest. In studies

involving a disease, a group of patients suffering from the disease is selected along with a matching group of healthy control subjects with similar background. The patient groups are sometimes further divided into separate sub-groups for patients at different stages of the disease or for those undergoing different types of treatment. The composition of the groups in shape characterization studies should reflect the intra-population variability due to age, sex, ethnicity, and other factors. It is often difficult and expensive to find a large number of qualifying individuals, so shape characterization methods are forced to work with small sample sizes.

Shape characterization methodology is a conduit between medical images and standard statistical methods. The latter require the information about shape to be represented by a fixed set of random variables, called *features*. Each individual in a sample must be represented by a list of numbers, which are the realizations of the features for that particular individual. The challenge of shape characterization is to derive an appropriate set of features to represent the relevant shape-related information contained in a collection of medical images and to describe each image as a list of feature realizations.

The transition from images to features draws on many areas of image analysis. Automatic segmentation algorithms may be used to find and extract objects of interest in medical images. The extracted objects need to be represented in a way that reflects their geometric properties. For example, it is common to represent an object as a mesh of points on the boundary. The structure of the mesh should be the same for all instances of the object in the sample, and corresponding mesh points in different instances should be placed in corresponding locations on the object. While it is possible to use all of the parameters that define an object representation as features, such a choice usually leads to poor statistical performance. It is usually advantageous to perform additional processing and filtering to derive features from the object representation.

Shape characterization research makes use of standard statistical techniques, such as *classification* and *density estimation*. Classification is used in applications that study the

effects of diseases on anatomical shape, and is at the center of shape-based diagnostic applications. Density estimation is used for tasks such as atlas building and automatic segmentation, where it is necessary to assess the validity of an object by assigning it a probability density score.

The tasks of classification and density estimation both consist of four phases, which are *model selection*, *training*, *testing*, and *application*. The model selection phase involves choosing an appropriate statistical model for the task at hand. For the task of density estimation, model selection involves choosing a probability distribution that can reasonably describe the variability present in the data. In classification, model selection means choosing one of many competing classifier methods. Statistical models are often selected empirically, using some heuristic knowledge about the nature of the problem. During the training phase, the parameters of the statistical model are *learned* using a training sample. In classification, the training sample consists of multiple instances from each of the classes, with class membership of each instance known. For density estimation, the training set simply contains multiple valid instances. During testing, the quality of the learned statistical model is evaluated by applying it to another sample, which is similar to the training sample, but is distinct from it. Testing can be used not only to validate the statistical model, but to re-learn its parameters. Finally, the application stage involves applying the statistical model to new data instances, for which the validity or class membership is not known.

A well selected set of features is of critical importance for classification and density estimation. For instance, a good choice of features may lead to the Gaussian distribution being appropriate for describing the variability in a sample, while a bad choice of features may lead to asymmetrical or multi-modal data. In classification, a good set of features includes those features that reflect the differences between classes, while excluding the features that capture intra-population variability and noise. A good set of features leads to classifiers that perform well during testing and application.

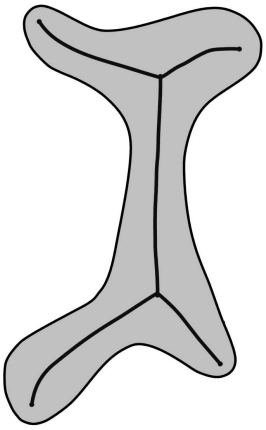


FIGURE 1.1. The medial locus of a two-dimensional object.

The overall goal of this work is to develop a general methodology for deriving effective features for various anatomical shape characterization problems. Contributions are made to two phases of the feature derivation process. One part of the dissertation deals with using the medial representation of object geometry as a means of producing "good" features. The other part of the dissertation deals with selecting the most relevant subset in a set of features derived from any geometric object representation. The following two sub-sections briefly motivate these two directions of research.

1.1.2. Medial Representation of Object Geometry. The medial loci of an objects, also known as skeletons, are a geometric construct widely used in computer vision and image analysis. One of several possible ways to define the medial loci is to use the grassfire analogy. In two dimensions, one pretends that the object is a patch of dry grass whose entire boundary is instantaneously set on fire. The locus of points where different fire fronts meet and extinguish themselves is the object's medial locus.

The medial locus of a two dimensional object, as shown in Fig. 1.1, consists of connected curve segments. Each segment corresponds to a part of the object, called *figure*, whose boundary has at most two positive maxima of curvature. Each point on the medial locus is expressed in terms of its position, as well at the time at which the

fire fronts reached that point. The latter measurement describes the local thickness of the object, i.e., the distance to the closest points on the boundary. The bending of the medial locus is indicative of the bending of the figures themselves, especially in the case of worm-like objects.

The power of medial loci lies in their ability to describe the geometric form of an object in terms of bending, thickness, and hierarchical composition into figures. For an anatomical object, its bending, thickness and figural composition may directly reflect the biological and physical processes that have formed and deformed the object. Therefore, the measurements derived from the medial loci are a promising basis for shape characterization features.

Traditional medial approaches derive the medial locus from the discrete representation of an object's boundary. Discrete boundary representations yield bushy medial loci with a very complex branching structure, in some cases creating nearly as many different branches as there are points describing the boundary. The literature contains a number of skeleton regularization methods, which eliminate the unstable and spurious parts of bushy medial loci, leaving only those branches that are relevant for qualitative shape description [Kimia et al. 1995, Siddiqi et al. 1999b, Ogniewicz and Kübler 1995, Näf et al. 1996].

These methods, however, are poorly suited for shape characterization, because they have no way of enforcing consistency of the branching topology of skeletons derived from a sample of similarly shaped objects. Even objects that are nearly identical may yield different medial loci using the boundary-to-medial approach. Since shape characterization requires all instances in a sample to be described using a fixed set of features, being able to guarantee a common medial branching topology is essential.

Using a representation of the skeleton computed by pruning Voronoi diagrams, Styner and Gerig [2001b], Styner [2001] have shown that such a common topology can be found

for populations of various anatomical objects. Styner and Gerig demonstrated that subcortical structures such as the lateral ventricle, the hippocampus, and the pallide globe, can be accurately described using a single medial sheet.

The ability to describe populations of objects using a fixed medial branching topology makes it possible to combine the geometrical expressive ability of medial loci with the theoretical principles of the deformable templates theory [Mumford 1996]. Such is the approach taken by the *m-rep* methodology. M-reps, short for medial representations, are deformable models that simultaneously describe the medial loci of objects and their boundaries. A single m-rep template with a fixed medial branching topology can be fitted to different instances of similar objects. M-reps have been shown by [Styner 2001] to be effective in accurately modelling populations of biological objects. While an m-rep does not give the precise medial locus of an object, it nevertheless captures its important qualitative geometric properties, such as bending, thickness and figural composition.

M-reps explicitly define the level of detail, or *scale*, at which they approximate objects. By varying the scale parameter, it is possible to describe a single object using a coarse-to-fine family of m-reps, each describing different qualitative aspects of the object's geometry.

This dissertation develops a framework for deriving features from coarse-to-fine families of m-reps and for leveraging the special properties of these features in order to localize shape variability. In the process, it introduces a novel new kind of m-reps that are particularly well suited for shape characterization.

1.1.3. Feature Selection. Classical statistical discrimination techniques are designed for problems with large samples and relatively small numbers of features. In high dimensional, low sample size problems commonly encountered in shape classification, such techniques tend to grossly over-fit the training data, resulting in poor generalization performance.

This problem has been addressed in the machine learning literature by a technique called *feature selection*. Feature selection is used to reduce the dimensionality of classification problems by finding the subset of features that best captures the differences between classes. Classifiers restricted to the selected subset of features are less affected by sampling noise and tend to generalize better than the classifiers trained on the entire feature set. Feature selection has been shown to dramatically improve the generalization ability of classifiers in high-dimensional problems [Bradley and Mangasarian 1998, Weston et al. 2001].

The potential benefit of using feature selection algorithms in shape characterization problems extends beyond the improved generalization ability. Examination of features deemed most relevant by such algorithms may reveal the areas of organs that are most affected by a disease, leading to improved localization and understanding of the biological processes responsible for the disease.

Feature selection algorithms in machine learning literature usually address general classification problems and make minimal assumptions as to where the different features come from and how they may be related. In shape classification problems, where features are usually derived from dense geometric object representations, there exist special relationships between neighboring features. By incorporating the knowledge of these relationships into feature selection algorithms, it is possible to improve their performance and stability when applied to shape classification. The two properties of shape features that are particularly useful for improving feature selection are thus *structure* and *locality*.

I use the term *structure* to indicate the importance of the order in which the features are arranged in a classification problem. In many problems the order of the features is arbitrary, as is the case, for example, when all the features describe different physical properties of an object, such as its height, weight, age or density. However, when the

features are measurements regularly sampled from a lattice, as is the case in many geometric object representations, the order of the features is important, as nearby features are more likely to be correlated than the far-away features.

A biological process responsible for variability in the shape of an anatomical object exhibits *locality* if it affects the object at one or at most a few locations, which are consistent across the population of objects. In reference to a feature set, I use the term locality to mean that some components of the statistical variability in the data can be localized to one or more subsets the features.

In the absence of structure and locality, the feature selection problem is purely combinatorial, since in the set of n features there are 2^n possible subsets and all of them are considered *a priori* to be equally worthy candidates for feature selection. The properties of structure and locality constitute prior knowledge about the kinds of feature subsets that ought to be selected. Feature sets consisting of one or a few contiguous subsets are more likely candidates than feature sets in which the selected features appear scattered. By assuming that shape features exhibit structure and locality, it is possible to effectively reduce the number of possible solutions of a feature selection algorithm.

This dissertation presents a novel feature selection algorithm that rewards the locality and structure of selected feature sets, and effectively improves the generalization rate of shape-based classifiers. The algorithm is applicable to features derived from various object representations.

1.2. Thesis

In this dissertation I assert and verify that

(1) *The multi-scale medial representation of object geometry can yield statistical features that (i) are appropriate and effective for the shape characterization tasks of classification and density estimation, (ii) can describe shape variability in*

intuitive terms of bending and thickness, and (iii) can detect and pinpoint local components of shape variability.

(2) *The localization and discrimination abilities of shape classification methods can be further improved by applying a feature selection algorithm that incorporates heuristics that reward locality of the selected set of features.*

(3) *The effectiveness of shape characterization based on the medial representation can be further improved by using a continuous medial representation.*

1.3. Claims

The following are the key claims and contributions made by this dissertation in the fields of anatomical shape characterization, medial object representation, and feature selection.

Claim 1. *New concept: A focus on locality in shape characterization.*

I assert that shape classification methods should make use of special properties of biological processes that are responsible for the variability in anatomical objects. In particular, I make an assumption that such processes exhibit *locality*, meaning that they affect anatomical objects at one or at most a few locations, which are consistent across the entire population. I then adapt shape characterization methodology to make use of locality.

Claim 2. *New method: Using a multi-scale medial object representation for localizing and describing components of shape variability.*

I develop a new shape characterization method that uses the medial representation of object geometry in order to describe shape differences intuitively, in terms of bending and growth. By using a medial representation that describes objects at different levels of detail, the method is able to find and describe local components of shape variability.

CLAIM 3. *New method: Selecting statistically relevant features in shape classification problems.*

I develop a new feature selection method that is tailored to shape classification problems. The method is unique because it looks for a set of relevant features that is not only small, but also structured and localized.

CLAIM 4. *Result: New feature selection method is shown superior to a similar method in synthetic examples.*

I generate a number of synthetic shape classification problems and show that the new feature selection method that rewards structure and locality of the selected feature set outperforms a similar feature selection method that does not make use of feature neighbor relationships. The simulations include two-dimensional examples based on the corpus callosum and three-dimensional examples based on the hippocampus.

CLAIM 5. *Application: New feature selection method is applied to clinical 3D hippocampus data.*

I apply the new feature selection method to real clinical data from a shape classification study that examines the relationship between schizophrenia and the shape of the hippocampus. I contrast my results with reported findings of the study.

CLAIM 6. *New method: Representing 2D and 3D object geometry using a continuous medial representation based on splines.*

I develop a new representation for object geometry that simultaneously defines the boundary and the medial locus of an object as functions on a continuous parameter space. The differential geometric properties of the medial locus and the boundary can be easily and accurately derived from the representation. The representation provides a powerful framework for establishing an object-centric coordinate system.

CLAIM 7. *Result: Continuous medial representation shown capable of describing and segmenting objects in 2D and 3D medical images.*

I apply the deformable segmentation technique to accurately fit a continuous medial template that has a complex branching structure to the cross-section of the vertebra in a 2D slice of a computer tomography (CT) image. I also show that the continuous medial representation can be automatically fitted to binary three-dimensional images of the hippocampus with a high level of accuracy.

1.4. Overview of the Chapters

The dissertation is organized into six chapters:

Chapter 2 presents the background information on the related methodology in image analysis and computer vision. The represented methods fall into areas of shape categorization, medial object representation and feature selection.

Chapter 3 presents the general framework for applying statistical shape characterization methodology to objects that are represented using coarse-to-fine families of discrete m-reps.

Chapter 4 presents the novel feature selection methodology that rewards structure and locality of the selected feature sets. It describes the algorithm and validates it using simulated data examples. The feature selection algorithm is applied to clinical hippocampus data.

Chapter 5 presents the novel medial representation, called continuous m-reps, which is constructed using splines. The chapter covers topics in 2D and 3D medial geometry, defines the representation, and shows results of using continuous m-reps for 2D image segmentation and 3D binary image segmentation.

Chapter 6 revisits the claims made in this thesis and discusses the work that remains to be done in order to combine the separate research accomplishments reported in this dissertation into a single cohesive shape characterization framework.

CHAPTER 2

Background

> In so far as the statements of geometry speak about reality, they are not certain, and in so far as they are certain, they do not speak about reality.
>
> Albert Einstein

2.1. Medial Representation of Geometrical Objects

This section presents the background material necessary to fully introduce and contextualize *m-reps*, the medial object representation that is central to this dissertation. The section is organized as follows. Subsection 2.1.1 defines the *medial locus*, which is a geometric construct widely used in computer vision to capture local symmetric properties of objects. Subsection 2.1.2 describes the structure of medial loci and their use for decomposing objects into simple figures. Subsection 2.1.3 explains the local geometric properties of medial loci. Subsection 2.1.4 presents a number of skeletonization algorithms used to extract medial loci from boundaries of objects. Subsection 2.1.5 defines medial atoms, which are the building blocks of m-reps and describes their use in core-tracking methods, which are a predecessor to m-reps. Subsection 2.1.6 talks about the theory of deformable models, which gives a foundation for the m-rep approach. Finally, subsection 2.1.7 describes the m-rep methodology.

The conceptual structure of this section is designed to introduce the reader to medial geometry first, then to show the traditional approach to medial representation of shape, which computes medial loci as a function of a given object, and finally to describe the

m-rep approach, which starts with a model that is medial in nature and fits the model to different instances of an object. After reading this section, the reader should understand the underlying medial geometry, which is modelled by m-reps, as well as the advantages of using m-reps for shape characterization purposes.

2.1.1. The Definition of the Medial Locus. Medial loci enjoy wide use in computer vision and image analysis, as well as in other fields of computer science such as graphics, computer aided design, and human-computer interfaces [Bloomenthal and Shoemake 1991, Sherstyuk 1999, Storti et al. 1997, Blanding et al. 1999, Igarashi et al. 1999]. The modern interest in medial loci originates with the work of Blum, who defined the medial locus of a two-dimensional object, studied its geometric properties, and noted its usefulness for describing the shape of objects [Blum 1967, Blum and Nagel 1978]. While the definition itself is deceptively simple, the thorough understanding of the properties of the medial locus requires rigorous mathematical treatment. Only recently have a number of mathematically rigorous studies of the medial loci of higher-dimensional objects been published [Damon 2002, Giblin and Kimia 2000].

I begin by defining the medial locus using a formulation that slightly extends Blum's original definition. Later in this section I mention some alternative definitions of the medial locus. The basic element of Blum's definition is the *maximal inscribed ball*.

Definition 1. *Let S be a connected closed set in $\mathbb{R}^n$. A closed ball $B \subset \mathbb{R}^n$ is called a* maximal inscribed ball *in S if $B \subset S$ and there does not exist another ball $B' \neq B$ such that $B \subset B' \subset S$.*

Next let us give a formal definition of the term *object* to avoid any ambiguity that the use of this word may bring.

Definition 2. *A set in $\mathbb{R}^n$ is called an n-dimensional* object *if it is homeomorphic to the n-dimensional closed ball.*

Let Ω denote an n-dimensional object and let $\partial\Omega$ denote its boundary. We are now ready to define the medial locus of Ω.

DEFINITION 3. *The* internal medial locus *of* Ω *is the set of centers and radii of all the maximal inscribed balls in* Ω.

DEFINITION 4. *The* external medial locus *of* Ω *is the set of centers and radii of all the maximal inscribed balls in the closure of* $\mathbb{R}^n/\Omega$.

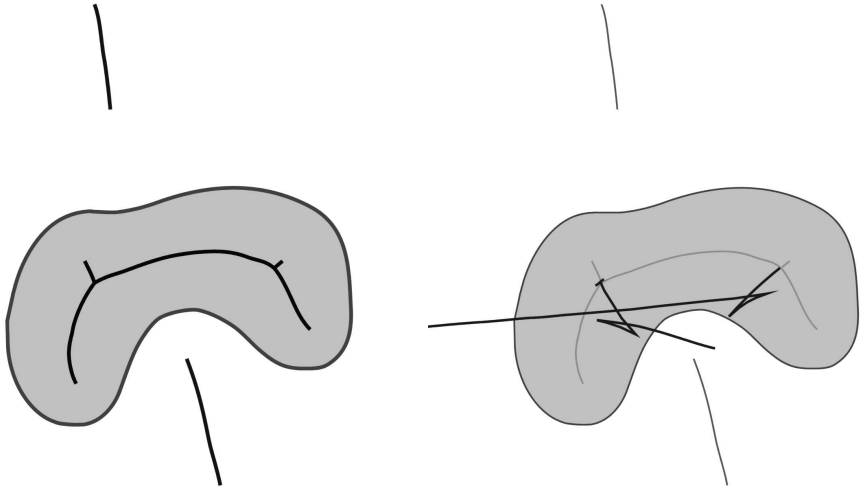
DEFINITION 5. *The* medial locus *of* Ω *is the union of its internal and external medial loci.*

DEFINITION 6. *The tuple* $\{\mathbf{x}, r\}$ *that belongs to the medial locus of an object* Ω *is called a* medial point *of* Ω.

The medial locus is thus a subset of the space $\mathbb{R}^n \times [0, +\infty]$. I will sometimes use the term medial locus to refer just to the set of the centers of the maximal inscribed balls, forgetting about their radii. The terms *skeleton*, *medial axis*, and *symmetric axis* have been used by other authors to describe both the internal medial locus and the entire medial locus as a whole, with or without the inclusion of the radial component.

It turns out that the medial locus consists of a countable number of manifolds whose codimension in the space $\mathbb{R}^n \times [0, +\infty]$ is no less than 2. Hence, the medial locus of a two-dimensional object consists of a number of curves and isolated points, and the medial locus of a three-dimensional object consists of surface patches, curves, and isolated points. The manifolds composing the internal medial locus lie inside the object and are bounded, while the manifolds in the external medial locus lie outside of the object and extend to infinity. Fig. 2.1a. shows the internal and external medial loci of a two-dimensional object.

Moreover, it turns out that the inscribed disks whose centers and radii compose the medial locus of an object generically are bitangent to the object's boundary. In fact,



a. b.

FIGURE 2.1. a. The internal and external medial loci of an object. b. The symmetry set of the same object, which contains the internal and external medial loci (shown in grey) as well as some cusped structures (black).

the medial locus is a subset of a more general geometric construct called the *symmetry set*, defined as the closure of the locus of centers and radii of all balls bitangent to the boundary of an object [Giblin and Brassett 1985]. The balls that generate the symmetry set are not restricted to lie either inside or outside of the boundary of the object. Hence, in addition to the medial locus, the symmetry set of an object contains connective structures such as the cusps shown on Fig. 2.1b.

A number of alternative geometric definitions of the medial locus of a planar object have been proposed in the literature. Leyton [1987] compares different definitions that start with maximal inscribed disks and construct the medial locus not only from the centers the disks but form other points as well. For example, Leyton's own definition (Process Inferring Symmetric Axis or PISA) uses the midpoints of the arcs connecting the points of bitangency between the inscribed maximal disks and the object boundary. A similar definition due to Brady uses the midpoints of the chords connecting the points of bitangency [Brady and Asada 1984]. The endpoints of the medial loci that result

from both of these definitions lie on the boundary of the object. In this dissertation I only work with Blum's definition of the medial locus and its extension to three or more dimensions.

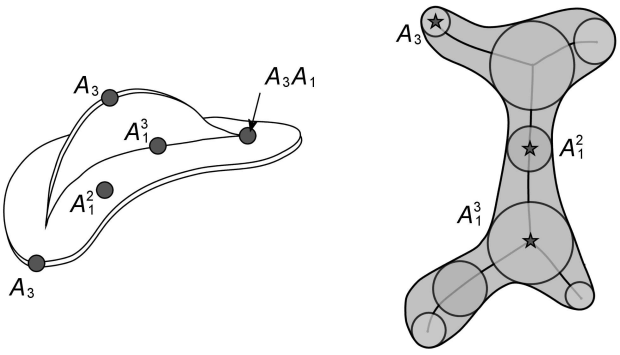
The medial locus can also be defined analytically using the following *grassfire analogy*. The object is imagined to be a patch of grass whose boundary is set on fire instantaneously. As the grass burns away, the fire fronts propagate inward and outward from the boundary. Grassfire propagation can be described by the following differential equation:

$$\frac{\partial \mathcal{C}(t,p)}{\partial t} = -\alpha \mathbf{N}(p), \tag{1}$$

where $\mathcal{C}(t,p)$ denotes the fire front at time t, parameterized by p, $\mathbf{N}(p)$ is the unit outward normal to the fire front, and α is a constant, positive for inward propagation and negative for outward propagation. As the propagation progresses, segments of fire fronts that originate from disjoint parts of the boundary begin to meet and quench themselves at points that are called *shocks*. The medial locus is defined as the set of all the shocks, along with associated values of time t at which each shock is formed. This analytic definition of the medial locus is equivalent to the geometric definition given previously; a proof was given by Calabi [1965a] and Calabi [1965b], Calabi and Hartnett [1968].

2.1.2. Structural Geometry of Medial Loci. Giblin and Kimia [2000] give a rigorous description of the structural composition and local geometric properties of medial loci of three-dimensional objects. Their description classifies medial points based on the multiplicity and order of contact that occurs between the boundary of an object and the maximal inscribed ball centered at a medial point.

Each medial point $P = \{\mathbf{x}, r\}$ in the object Ω is assigned a label of form A_k^m. The superscript m indicates the number of distinct points at which a ball of radius r centered at $\mathbf{x}$ has contact with the boundary $\partial\Omega$. The subscript k indicates the order of contact between the ball and the boundary.



a. b.

FIGURE 2.2. a. Different classes of points that compose the medial locus of a three-dimensional object, as categorized by Giblin and Kimia. b. Three possible ways in which maximal inscribed disks can be tangent to the boundary of a two-dimensional object.

The order of tangent contact is a number that indicates how tightly a ball B is fitted to a surface S at a point of contact P, and can take take the following values:

A_1 **contact:** B is tangent to S at P;

A_2 **contact:** B is one of the spheres of curvature of S at P but not at a ridge [1] of the corresponding curvature. In other words, B is tangent to S at P and one of the principal curvatures of S at P, say κ_i, equals the reciprocal of the radius of B. However, κ_i does not attain a local extremum at P;

A_3 **contact:** P lies on a ridge of S, and B is a sphere of curvature of S at P;

A_4 **contact:** P is a turning point on a ridge of S.

The following theorem specifies all the possible types of contact that can generically occur between the boundary of a three-dimensional object and the maximal inscribed balls that form its medial locus. The theorem also specifies how medial points with different associated type of contact are organized to form surfaces and curves in the medial locus.

[1]The term *ridge* has multiple uses. In this context it refers to the locus of points on a surface at which one of the principal curvatures is positive and has a local maximum or is negative and has a local minimum.

THEOREM 1 (Giblin and Kimia). *The internal medial locus of a three-dimensional object Ω generically consists of*

(1) *sheets (manifolds with boundary) of A_1^2 medial points;*
(2) *curves of A_1^3 points, along which these sheets join, three at a time;*
(3) *curves of A_3 points, which bound the free (unconnected) edges of the sheets;*
(4) *points of type A_1^4, which occur when four A_1^3 curves meet;*
(5) *points of type A_1A_3 (i.e., A_1 contact and A_3 contact at a distinct pair of points) which occur when an A_3 curve meets an A_1^3 curve.*

Proof. See [Giblin and Kimia 2000] for a rigorous proof. □

In two dimensions, a similar classification of medial points is possible. The internal medial locus of a two-dimensional object generically consists of (i) curves of bitangent A_1^2 points, (ii) points of type A_1^3 at which these curves meet, three at a time, and (iii) points of type A_3 which form the free ends of the curves. The three classes of contact are illustrated in Fig. 2.2a. In two dimensions, A_3 contact means that the inscribed disk and the boundary osculate at a local maximum of boundary curvature.

The geometric properties of the external medial locus are similar to those of the internal locus, with the exception that the sheets and curves are no longer completely bounded and may stretch out to infinity. Less effort has been devoted in the literature to the study of external medial loci.

Theorem 1 states that each surface composing the internal medial locus of an object joins another two such surfaces or terminates at a point of type A_3^1, which correspond to ridges of curvature on the boundary surface. Similarly, curve segments composing the internal medial loci of two-dimensional objects either connect with pairs of other curve segments or terminate at points corresponding to positive maxima of boundary curvature. Hence, the number of such ridges or maximal points limits the number of surfaces and curves in the internal medial locus. It can be shown by induction that the number of

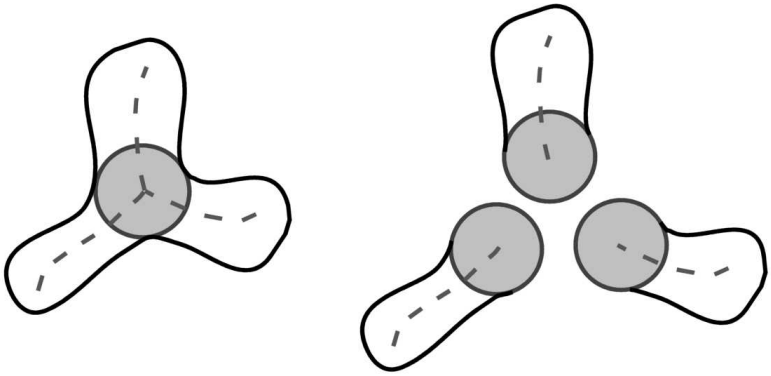


FIGURE 2.3. Decomposition of a planar object into figures with joints. Each curve in the medial locus corresponds to a single figure.

curve segments composing the internal medial locus of an object whose boundary has M positive maxima of curvature may not exceed $2M - 3$.

I will use the term *stratum* to refer collectively to curves in medial loci of two-dimensional objects and to surfaces in medial loci of three dimensional objects. The composition of medial loci into interconnected strata makes it is possible to decompose geometrically complex objects into simple components called *figures*. Roughly speaking, a figure is the part of an object that corresponds to a particular stratum in the medial locus. A particularly simple mathematical definition of a figure is the following:

DEFINITION 7. *The union of closed balls whose centers and radii form a single stratum in the medial locus of an object is called a* ***figure with joint***[2].

Figures generated by strata belonging to the internal medial locus of an object are bounded, and the union of all such figures is the object itself. The intersection of a pair of figures with joints is non-empty if the generating strata of the two figures are connected. This non-empty intersection is called the *joint*, and it comprises of balls of triple boundary contact.

[2]As distinguished from *figure*, which will be discussed later in the context of m-reps

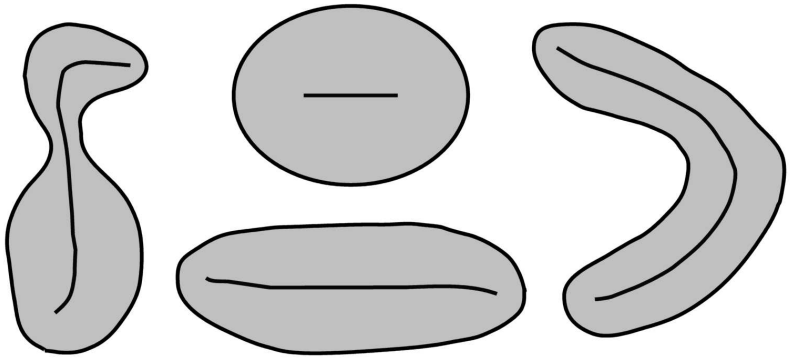


FIGURE 2.4. All four of these objects fall into the category of *figures with joints* according to Def. 7, even though none of them have an actual "joint". Notice that the figure on the right has more than two positive maxima of curvature.

The internal medial locus of a figure has only a single stratum and figures can be said to be geometrically simple and easier to study than whole objects. Fig. 2.4 shows some examples of two-dimensional figures with joints.

The relationship between the structure of symmetry sets (which are a superset of medial loci) and the extrema of boundary curvature of two-dimensional objects are central to Leyton's theory of symmetry [Leyton 1987]. For planar objects, Leyton's curvature-symmetry duality theorem states that

> Any section of curve, that has one and only one curvature extremum, has one and only one symmetry axis. This axis is forced to terminate at the extremum itself.

The extension of this theorem to three dimensions is given by Yuille and Leyton [1990].

Leyton's theory states that the curves composing the symmetry set of an object represent the history of events that have formed the object. According to Leyton's postulates, "memory is always in the form of asymmetry," meaning that asymmetry makes it possible to recover information about the formation of an object, while "symmetry is always the absence of memory." The more complex the structure of an object's symmetry set,

the more asymmetry is there in the object, the more can we learn about its formation. Moreover, Leyton states that the extrema of curvature are the places where the boundary has been pushed in from the outside, or pushed out from the inside, indicating growth. The medial curves that terminate at these extrema are in a sense arrows in the direction of the push. Hence, the symmetry set is a diagram of protrusion and indentation operations that have been applied to an object [Leyton 1992].

Other researchers have shown that figural decomposition of objects often corresponds to the cognitive processing performed by the human brain. Quoting Chapter 2 of Katz [2002],

> Burbeck and Pizer [1995]... show how figures encompass many of the phenomena shown in psychophysical studies. Figures are formed at the places of highest negative boundary curvature, which studies have shown are the places where we visually decompose objects [Hoffman and Richards 1984, Biederman 1987, Braunstein et al. 1989]. The junctions of figures also have special importance, matching Biederman's work on the junctions of visual parts [Biederman 1987]. The ends of figures have also been shown to match work showing their special visual significance [Hubel and Wiesel 1977, Orban et al. 1979, Leyton 1992]. [Rock and Linnett 1993] shows that figures are captured preattentively by our visual systems.

2.1.3. Local Geometry of Medial Loci. Prior to describing the local geometry of medial loci, let us introduce a useful notation for referring to the points of contact between a ball places at a medial point and the boundary of an object.

DEFINITION 8. *If $P = \{\mathbf{x}, r\}$ is a medial point of an object Ω, then the set of points of contact between a ball of radius r centered at $\mathbf{x}$ and $\partial\Omega$ is called the* boundary pre-image *of P.*

In other words, a medial point labelled A_k^m has a boundary pre-image that contains m points. Since at most of the medial points m is equal to 2, the following definition is quite useful.

DEFINITION 9. *If points A and B form the boundary pre-image of a medial point P, then A is called a* medial involute *of B and vice versa.*

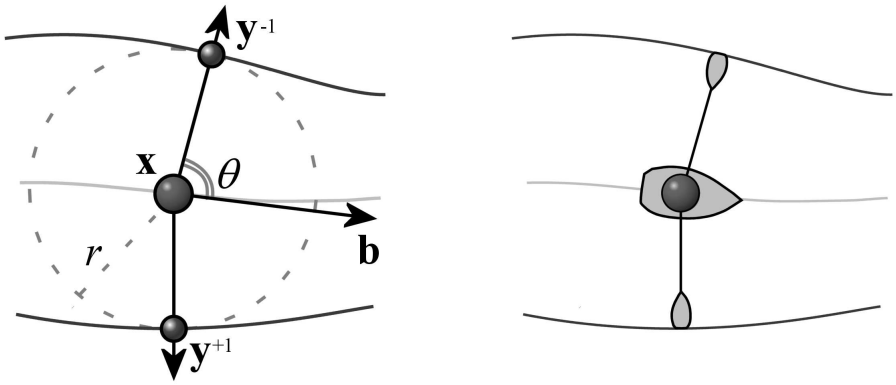
Said in another way, medial involutes are pairs of points on the boundary of an object that are symmetric with respect to the medial locus. It is possible for a point to have multiple medial involutes, for example one with respect to the interior medial locus, and one with respect to the exterior medial locus.

The major part of Blum's work on the internal medial loci of two-dimensional objects is devoted to the study of the geometric relationships between medial points and their boundary pre-images [Blum 1967, Blum and Nagel 1978]. Blum showed that the points in the boundary pre-image can be expressed in terms of the position and radius of the medial point and from their derivatives with respect to movement along the medial locus.

For the purposes of studying local geometry of internal two-dimensional medial loci it suffices to focus on medial points that lie on interior of the curves composing the medial locus and thus have two-point boundary pre-images. The geometric properties of the free endpoints and connecting endpoints of medial curves can be derived as the limit cases of the interior point properties.

In addition to using the position $\mathbf{x}$ and the radius r to characterize each point on the medial locus, Blum uses two first order properties. The first property is the slope of the medial curve at the medial point, which can be expressed as an angle α, a unit-length tangent vector $\mathbf{b}$, or as a rotation matrix $\mathbf{R}$. The second is called the *object angle* and is given by

$$\theta = \arccos\left(-\frac{dr}{ds}\right), \tag{2}$$



a. b.

FIGURE 2.5. Local medial geometry. a. Local geometric properties of a medial point and its boundary pre-image. b. The rowboat analogy for medial points.

where s is the arc length along the medial curve. The object angle is indicative of the narrowing rate of the object with respect to movement along the medial curve. When the object angle is equal to $\pi/2$, the radius has a critical point and as one moves along the medial locus, the object retains its local width.

Given a medial point characterized by $\mathbf{x}, r, \mathbf{R}$, and θ, the two points $\mathbf{y}^1$ and $\mathbf{y}^{-1}$ comprising its boundary pre-image are given by

$$\mathbf{U}^{\pm 1} = \mathbf{R} \begin{bmatrix} \cos(\theta) \\ \pm \sin(\theta) \end{bmatrix}, \tag{3}$$

$$\mathbf{y}^{\pm 1} = \mathbf{x} + r\mathbf{U}^{\pm 1},$$

where $\mathbf{U}^1$ and $\mathbf{U}^{-1}$ are unit-length vectors orthogonal to the boundary of the object at $\mathbf{y}^1$ and $\mathbf{y}^{-1}$.

Fig. 2.5a describes the local geometry of a medial point and associated boundary pre-image points $\mathbf{y}^{-1}$ and $\mathbf{y}^1$. The angle formed by the points $\mathbf{y}^{-1}$, $\mathbf{x}$, and $\mathbf{y}^1$ is bisected by the vector $\mathbf{b}$, the unit tangent vector of the medial curve at $\mathbf{x}$. The angle between $\mathbf{b}$ and the vectors $\mathbf{y}^{-1} - \mathbf{x}$ and $\mathbf{y}^1 - \mathbf{x}$ is the *object angle*.

The quantities $\mathbf{x}, \mathbf{y}^{\pm 1}, r, \mathbf{b}$, and θ appear frequently in this dissertation. To better remember these quantities, consider an analogy between a medial point and a one-person rowboat, illustrated in Fig2.5b. The position of the rower in the boat corresponds to $\mathbf{x}$, and the length of the oars corresponds to r. The vector $\mathbf{b}$ represents the direction in which the boat is moving and θ is the angle that each oar makes with $\mathbf{b}$. The points $\mathbf{y}^{-1}$ and $\mathbf{y}^{1}$ correspond to the tips of the oars, and the directions of the oars are given by the vectors $\mathbf{U}^{1}$ and $\mathbf{U}^{-1}$. The movement of a point along the medial locus is analogous to the rowboat navigating down the middle of a stream, with the rower adjusting his oars in such a way that their tips always just touch the banks of the stream (of course, the oars are made of a stretchable material, and as the boat moves, their length changes). A similar analogy to a sailboat is made in m-rep literature, and the term *sail vector* is used instead of the term *oar*.

The values of $\mathbf{x}$, r, and their derivatives can be used to qualitatively describe the local bending and thickness of an object. The measurements $\mathbf{x}$ and $\mathbf{b}$ along with the curvature of the medial curve describe the local shape of the medial locus, and subsequently describe how a figure bends at $\mathbf{x}$. A figure that has a line for its medial curve is symmetrical under reflection across that line. The measurement r describes how thick the figure is locally, while $\cos\theta$ describes how quickly the object is narrowing with respect to movement along the medial curve. A figure with a constant value of r has the shape of a worm.

Free ends of medial curves, where the maximal inscribed disk and the boundary osculate and the boundary pre-image contains a single point, are a limiting case of the bitangent disk situation. As our imaginary rowboat approaches such a point, its oars come closer and closer together until they collapse infinitely quickly at the end-point, forming a single vector in the direction $\mathbf{b}$. The object angle θ, which is equal to half of the angle between the oars, is zero at such endpoints.

The geometry of medial loci of three-dimensional objects is considerably harder to visualize and express than the planar medial geometry. A number of researchers have

studied the differential geometry of three dimensional medial loci [Nackman 1982, Vermeer 1994, Gelston and Dutta 1995, Hoffmann and Vermeer 1996, Teixeira 1998].

Chapter 5 presents in detail the relationship between medial points of three-dimensional objects and the boundary pre-images of these points. Many of the results reported there are based Damon's work on *skeletal sets* [Damon 2002]. This work follows the generative approach to medial geometry, as opposed to the previously described approaches that derive the medial locus from the boundary description of a given object. In the generative approach, the medial locus is defined first, and the object and its boundary are generated by outward flow from the medial locus. As the following sections demonstrate, the generative approach to medial geometry is more applicable for problems of object modelling and shape description than the derivative approaches. The generative approach is the cornerstone of m-rep methodology.

Damon's skeletal set is a stratified set[3] of arbitrary dimension, on which a multi-valued vector field, called the *radial vector field* is defined. At most points in the skeletal set a pair of radial vectors is defined; these vectors point in the different directions relative to the tangent space of the skeletal set. At edges of skeletal manifolds (i.e., the boundaries of the manifolds with boundary) that are not shared by more than one manifold the radial vectors come together to form a single vector that lies in the tangent space of the manifold. At shared edges, more than two radial vectors are defined. The endpoints of the radial vectors form a locus that is called the boundary of the object described by the skeletal set. Damon describes a number of constraints that must be satisfied by the skeletal set and the radial vector field in order for the boundary to be continuous and differentiable. These constraints are expressed in terms of the *radial shape operator*, which measures how the radial vectors bend with respect to the skeletal set. This operator not only describes the local properties of the radial vector field but can also be used to express the local differential geometry of the boundary.

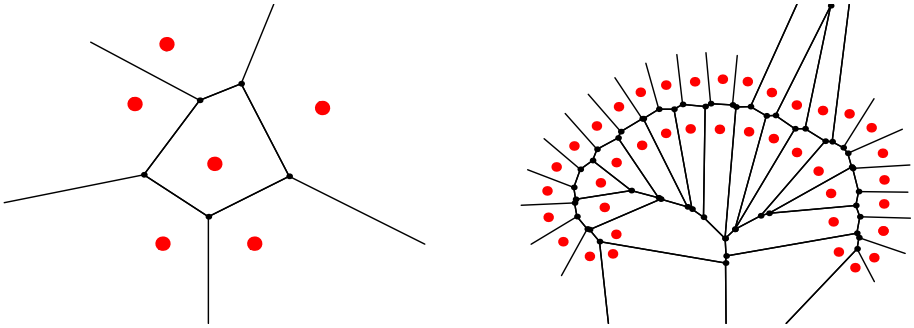
[3]A stratified set consists of interconnected *manifolds with boundary* of different codimension.

The Blum medial locus can be constructed as a special case of the general skeletal set by requiring that the radial vectors at each point be symmetric with respect to its tangent space of the skeletal set.

2.1.4. Extracting Medial Loci of Objects. The computer vision literature describes a large number of *skeletonization* methods, which extract medial loci of objects starting from some boundary representation. In most practical applications the object and its boundary are represented discretely, for example as a set of pixels of the same intensity in a characteristic image or as mesh of points. Skeletonization is made difficult by the inherent sensitivity of the medial locus to the fine details of the boundary representation. Given two different discrete representations of the same object, the true medial loci of the two representations can have a different *medial branching topology*, i.e., a different number of figures and a different connectivity graph between the figures.

Hence, the challenge of skeletonization is not to find the precise medial locus of an imprecisely specified boundary, but rather it is to compute an approximation of the medial locus that is consistent with respect to different discrete representations of the same object. Moreover, a good skeletonization method should yield similarly structured medial loci for objects that are similar objects and for versions of the same object that have been differently rotated and magnified.

This subsection offers a detailed look into two approaches to robust skeletonization: hierarchic Voronoi skeletons and shocks of boundary evolution. These two approaches, together with the core tracking approach described in Subsection 2.1.5, are compared in the overview paper by Pizer et al. [2002]. A special property of these approaches is that they provide a scale parameter that makes it possible to tune the accuracy with which the result matches the precise medial locus of the input boundary. The loci computed at larger values of the scale parameter generalize better to different discrete representations



a. b.

FIGURE 2.6. Examples of Voronoi diagrams. a. Voronoi Diagram of six points. b. Voronoi Diagram of points sampled from the boundary of the corpus callosum. The skeleton of the object is just the internal portion of the Voronoi Diagram.

of the input object as well as to other similar objects. For completeness, a number of alternative skeletonization methods are referenced at the end of this sub-section.

2.1.4.1. *Voronoi Skeletons.* Voronoi skeletons [Ogniewicz 1993, Székely 1996] are computed by calculating the *Voronoi diagram* of a set of points sampled from the boundary of an object. Fig. 2.6a shows a Voronoi diagram of a set of six points on a plane. The diagram consists of *Voronoi regions*, which are sets of points located closer to a particular generating point than to any other generating point. The line segments in the diagram are called *Voronoi edges*; they separate Voronoi regions and are loci of points that are equidistant from a a pair of generating points. The points where Voronoi edges meet are equidistant from three or more generating points.

When the generating points of a Voronoi diagram are sampled from the boundary of an object, as shown in Fig. 2.6b, the similarity between Voronoi edges and the curves composing the medial locus becomes apparent. A circle of appropriate radius centered at a point on a Voronoi edge contains two generating points, i.e., has two points of contact with the boundary. A circle centered at an intersection of two Voronoi edges contains three generating points, i.e., has three points of contact with the boundary, as do disks centered at intersections of curves in the medial locus. The Voronoi diagram is

also related to the grassfire analogy: if some points on the boundary are set on fire (as opposed to the whole boundary), the places where fire fronts meet and quench themselves are the edges in the Voronoi diagram of these points.

The Voronoi diagram of a set of boundary points contains edges that extend outside of the object, possibly to infinity. The discrete approximation of the boundary obtained by connecting the generating boundary points with line segments cuts the Voronoi diagram into internal and externals parts. The internal part is called the *Voronoi skeleton*.

The Voronoi skeleton generated by a discrete representation of an object's boundary is an approximation of that object's internal medial locus. Schmitt proves that as the number of generating boundary points increases, the Voronoi skeleton converges in the limit to the continuous medial locus, with the exception of the edges generated by neighboring pairs of boundary points [Schmitt 1989].

The Voronoi skeletons, such as the one shown in Fig. 2.6b, contain many branches, some of which are spurious and sensitive to the slightest changes to the generating boundary points. For instance, a Voronoi skeleton computed from the set of pixels forming the boundary of an object in an image can change significantly if the object in the image is rotated.

In order to make Voronoi skeletons more robust to small boundary changes, researchers have proposed to isolate parts of the Voronoi skeletons that are most stable and significant. A number of measures of significance for edges and groups of edges in the Voronoi skeleton have been introduced in the literature [Ogniewicz and Kübler 1995, Székely 1996]. The significance measures make it possible to establish trunk-branch relationships between connected edges in the Voronoi skeleton, and thus to establish a hierarchy of figures and sub-figures. The edges that fall far from the root of this hierarchy and have small significance values do not contribute to the descriptive ability of the Voronoi skeleton and are trimmed. Trimming on the basis of significance introduces a component of scale into Voronoi skeletonization. By adjusting the threshold level of

significance at which edges are discarded from the skeleton, it is possible to generate a coarse-to-fine spectrum of skeletons.

Both local and global measurements of significance have been proposed to organize and trim Voronoi skeletons. Local measurements assign a significance score to each edge in the Voronoi skeleton using a heuristic, such as the distance along the boundary between the pair of generating points to which the edge is equidistant [Ogniewicz 1993]. Global methods, on the other hand, compute the significance of an edge or a group of connected edges by measuring its impact on the appearance of the whole object, for example measuring the effect that removing the edge or edges from the skeleton would have on the shape of the boundary[Näf 1996, Styner 2001, Katz 2002].

The construction of Voronoi skeletons of three dimensional objects, while analogous to the two-dimensional construction, is much more difficult to implement. One difficulty arises during the traversal of Voronoi edges. The connectivity of Voronoi edges in two dimensions organizes them into a tree structure which can be traversed from the trunk to the leaves; in three dimensions the Voronoi edges may form a graph that contains cycles and is more difficult to traverse. Methods for organizing three-dimensional Voronoi skeletons into figures have been developed by Näf [1996], Attali et al. [1997], and Styner [2001].

2.1.4.2. *Shocks of boundary evolution.* As mentioned in subsection 2.1.1, the medial locus of an object can be defined as a set of points, called shocks, where a grassfire instantaneously started at the boundary of the object extinguishes itself. A number of methods for finding the medial locus by simulating the grassfire propagation equation has been developed [Kimia et al. 1995, Siddiqi et al. 1998, 1999a, August et al. 1999].

The methods of Kimia et al. [1995] and Siddiqi et al. [1998] start with a parameterized object boundary, $C(p)$, and evolve this boundary over time using the following differential

equation:

$$(4) \qquad \begin{cases} \frac{\partial C(t,p)}{\partial t} &= (\alpha - \beta\kappa)\mathbf{N} \\ C(0,p) &= C(p) \end{cases}$$

In this equation N denotes the unit outward normal to $C(p)$ at p, κ denotes the curvature of $C(p)$ at P, and α and β can be functions or simply constants. The deformation depends only on the local geometry of the boundary at time 0. The component $\beta\kappa\mathbf{N}$, known as *mean curvature flow*, has a regularizing effect on the deformation. When β is zero, the deformation is equivalent to the grassfire transform (1).

In cases where the curvature term of the curve evolution equation (4) is zero or is negative, the evolution results in formation of sharp corners and folds. However, since the evolving curve represents the boundary of an object and therefore may not cross itself, such folds must not be allowed to happen. According to Sethian [1996], the evolution of curves that enclose matter must adhere to the *entropy condition*, which requires that once the evolving curve forms a corner, matter is permanently dislodged from the curve, so as to prevent folding. The formation of a corner and the subsequent dislodgement of curve points is referred to as a *shock*. Sethian shows that the entropy condition poses a problem for any straightforward numerical approach to curve evolution and proposes the use of the *level set method*, which embed the boundary curve as a level set of a higher-dimensional manifold, to which evolution can be applied in a robust fashion [Sethian 1996].

Kimia et al. [1995] and Siddiqi et al. [1998] use level set methods to find the shocks of the evolution equation (4). An alternative approach to finding the shocks using distance transforms and the Hamilton-Jacobi equation was more recently introduced by Siddiqi et al. [1999b].

Kimia et al. classify the shocks of the boundary propagation equation into four types. A first order shock is a discontinuity in the normal of the boundary of the object, i.e., a

positive maximum of curvature that at some point during the evolution forms a corner. A second order shock is formed when two opposing boundaries collide at a single point. A third order shock occurs when the two boundaries collide not just at a point, but along a small neighborhood. Third order shocks occur when boundaries are parallel. Fourth order shocks are events when a closed boundary collapses to a single point. This classification is similar to the geometric classification of medial points due to Giblin and Kimia [2000], which is described in subsection 2.1.2.

Siddiqi et al. [1999a] gives a framework for describing planar objects using directed acyclic shock graphs. Each vertex in the graph represents a shock of type 2 and 4 or a connected set of shocks of type 1 and 3 since the latter do not occur at isolated points. Each shock or shock group is indexed by a number and by the time (or times) of formation. The structure of the graph is the reverse of the grassfire transform. At the root are the last shocks to be formed (which correspond to the most important shape features), and the leaves contain the earliest shocks, ones most sensitive to boundary noise. The graph is converted to a tree. Two shapes are compared using a recursive tree matching algorithm which matches vertices based both on the topology of their rooted subtrees and the types of shocks at each vertex. The algorithm successfully matches up instances of similar shapes even in presence of noise and occlusions.

By organizing shocks into a tree structure with nodes at different levels, Siddiqi et al. imposes a significance measure on the shocks. This significance measure can be used to extract the robust components of the medial locus of an object similarly to the way a significance measure is used to trim Voronoi skeletons. By selecting different thresholds of this significance measure, it is possible to produce a coarse-to-fine family of medial loci.

August et al. [1999] use the concept of ligature as a measure of significance of parts of the medial locus. Ligature is defined by Blum as a set of medial points whose boundary pre-image contains a concave corner of the boundary. August generalizes ligature

to medial points whose disks touch the boundary in an ϵ-neighborhood of a negative maximum of curvature. Regions of the medial locus are then characterized in terms of stability. The most stable are the order 4 shocks (also called 'seeds') and order 3 shocks generalized to almost parallel boundaries (also called 'bends'). The least stable segments correspond to the ligature.

August, et al. demonstrate that by removing ligature one can simplify the medial locus and remove unstable branches without changing the shape of the object significantly. The results are applied to shock matching: as ligature nodes are removed from shock graphs, similarly shaped objects become more likely to have the same shock graph. August, et al. also show that the instabilities associated with ligature can not simply be removed by smoothing the boundary with diffusion and hence require special treatment.

2.1.4.3. *Other skeletonization approaches.* In this subsection I presented two skeletonization methods that can produce a coarse-to-fine family of medial loci by adjusting the value a scale parameter. The ability to produce medial descriptions at different levels of detail is critical for a skeletonization method, as without it the method is prone to have stability issues. However, in addition to the two approaches just described, the computer vision literature proposes many other skeletonization algorithms. Among them are based on mathematical morphology, others based on constructive solid geometry and yet other based on Euclidean distance transforms.

The mathematical morphology approach [Serra 1982, Matheron 1998, Malandain et al. 1993, Jonker and Vossepoel 1995] uses as its input the binary characteristic image of an object. Erosion is applied to the binary image iteratively, thinning the object until it is one pixel thick. Much research has been focused on finding structural elements (kernels) that preserve the topology of the object through erosion. Erosion is inherently a discrete process and it is highly sensitive to the rasterization of the object. Medial loci yielded by erosion are sensitive to the rotation and magnification of the object prior to the imaging

process. Working at pixel level, erosion often can not discern the local geometry of the medial locus near branch points.

The medial locus of an object represented by a binary characteristic image also can be approximated as the set of ridges of the Euclidean distance transform [Danielsson 1980, Arcelli and Sanniti di Baja 1992]. For instance, Golland et al. [1999] obtain a medial description of an object by fitting a snake to the critical set of the distance transform. The element of scale can be introduced by blurring the distance transform image with a Gaussian kernel. Ends of distance transform ridges are difficult to find and the branching topology of these ridges is sensitive to the resolution of the input image. This sensitivity is exhibited not only at the level of branches that can be easily removed by pruning but also at branches nearer the root of the branching hierarchy.

The constructive solid geometry approach [Lee 1982, Culver 2000, Sherbrooke et al. 1996] is geared towards human-made objects with boundaries that have sharp corners and can be described accurately using a small number of polygons. This approach finds the exact medial locus of the polygonal representation of an object's boundary. Such a locus consists of polygons and paraboloid surface patches. Elements of this approach are similar to Voronoi skeletonization.

2.1.5. Medial Atoms and Core Tracking. I dedicate an entire section to core tracking because it is a skeletonization approach that was the precursor to the m-rep methodology, which is used extensively in this dissertation. In particular, the concept of medial atoms, which is central to m-reps, evolved from the core tracking research. A reader interested in understanding the upcoming chapters of this dissertation should become familiar with the contents of the first part of this section, which is devoted to medial atoms.

A medial atom is a modelling primitive that represents a place on the medial locus of an object. A medial atom describes such a place up to a specified differential order

and with a specified level of tolerance. The continuous medial locus of an object can be discretely sampled into a set of medial atoms. However, medial atoms should be thought of as entities that exist independently any medial locus, as demonstrated by the following definition, which makes no mention of the latter.

DEFINITION 10. *An* n-dimensional medial atom of order 0 *(*$n = 2, 3$*) is a tuple* $m = \{\mathbf{x}, r\}$ *that satisfies*

$$\mathbf{x} \in \mathbb{R}^n, \ r \in (0, +\infty). \tag{5}$$

Geometrically, a medial atom of order 0 is simply interpreted as a ball. Such medial atoms essentially correspond to maximal inscribed balls whose centers and radii form the medial loci of objects. Given a structured collection of medial atoms of order 0 sampled from the medial locus of an object, is is possible to approximately reconstruct the object's boundary by 'shrink-wrapping' the balls defined by the atoms, as shown in Fig. 2.7.

The use of the word "order" in the above definition refers to the fact that medial atoms can be used to approximate medial loci up to a given order. An order 0 atom describes zeroth order medial properties, which are position and radius. An order 1 atom, which is defined below, describes the derivatives of position and radius.

The shortcoming of medial atoms of order 0 and the shrink-wrap boundaries derived from them lies in the fact that while both the medial locus and the boundary can be approximately recovered from a set of order 0 medial atoms, the local symmetry relationships between pairs of medial involutes can not be directly reconstructed. However, medial atoms of higher order can be used to capture these symmetry relationships and are more useful for object representation.

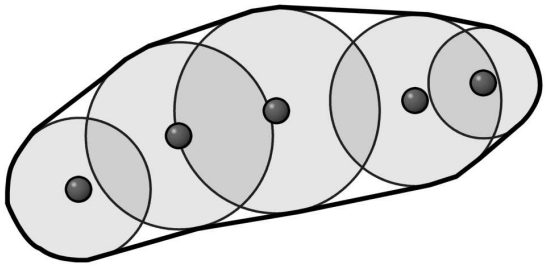


FIGURE 2.7. Boundary of an object reconstructed by shrink-wrapping a collection of order 0 medial atoms.

DEFINITION 11. *An* n-dimensional medial atom of order 1 *is a tuple* $m = \{\mathbf{x}, r, \mathbf{R}, \theta\}$ *that satisfies*

$$\mathbf{x} \in \mathbb{R}^n,\ r \in [0, +\infty),\ \mathbf{R} \in \mathrm{SO}(n),\ \theta \in [0, \pi/2]\ , \tag{6}$$

where $\mathrm{SO}(n)$ *is the symmetric orthogonal group, i.e., the group of all rotations in* n*-dimensional Euclidean space.*

The additional components of the medial atom of order 1 are $\mathbf{R}$, called its *orientation*, and θ, called its *object angle*. Recall that these are precisely the same first order quantities that were used in equation (3) to relate medial points to their boundary pre-images.

The orientation of an atom can be expressed as a rotation matrix $\mathbf{R}$, or equivalently as an orthogonal frame of unit vectors $\{\mathbf{f}_1 \ldots \mathbf{f}_n\}$ placed at the position $\mathbf{x}$. In two dimensions, $\mathbf{f}_1$ defines the tangent direction of the medial curve at $\mathbf{x}$ and $\mathbf{f}_2$ defines the normal direction. In three dimensions, the vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ define the tangent plane of the medial surface at $\mathbf{x}$, with $\mathbf{f}_1$ defining the direction of greatest decrease in r.

The rowboat analogy used to describe the first order geometry at a point on the medial locus can be used equally well to describe medial atoms of order 1. The atom's

position and radius correspond to the position of the rower and the length of the oars; the orientation corresponds to the direction of the boat's bow, and the object angle to the angle between the bow and the oars. Unlike their order 0 cousins, which one could visualize as an inflated rubber tube (a rather low-tech floatation device), medial atoms of order 1 explicitly define a pairs of points on the boundary of the object that they describe. These points are, of course, the tips of the oars and are given by the following definition.

DEFINITION 12. *The tuples* $\{\mathbf{y}^{-1}, \mathbf{U}^{-1}\}$ *and* $\{\mathbf{y}^{1}, \mathbf{U}^{1}\}$ *given by*

$$\mathbf{U}^p = \mathbf{R} \begin{bmatrix} \cos(\theta) \\ p\sin(\theta) \\ 0 \end{bmatrix}, \quad p = \pm 1 \ , \tag{7}$$

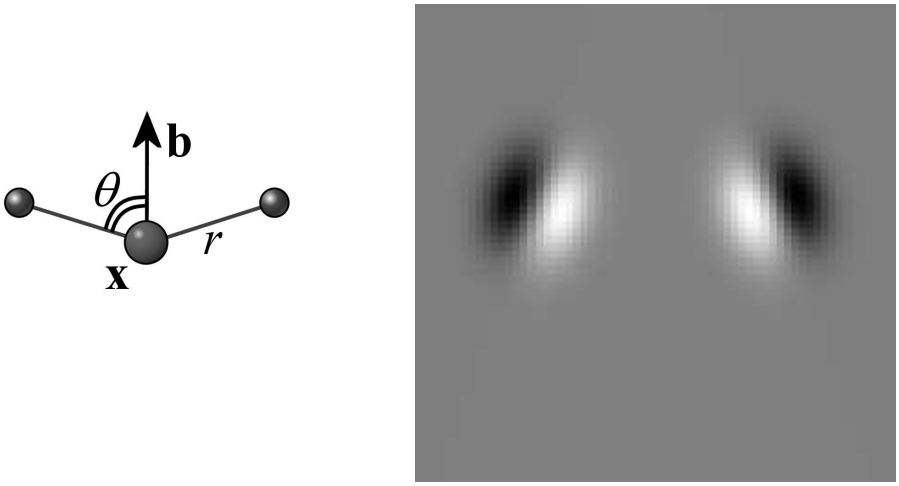
$$\mathbf{y}^p = \mathbf{x} + r\mathbf{U}^p$$

are called the implied boundary nodes *of the medial atom* $m = \{\mathbf{x}, \mathbf{r}, \mathbf{R}, \theta\}$.

Medial atoms of order 1 correspond to medial points of type A_1^2, whose boundary pre-images contain two points.[4] Medial points of other types can be represented either using special medial atoms, as discussed later in Sec. 2.1.7, or as limit cases of A_1^2 medial points. In this dissertation I deal only with order 1 medial atoms because these atoms contain all the information necessary to reconstruct the boundary pre-images of medial points and the symmetry relationships between pairs of medial involutes.

Medial atoms are used by *core tracking* methods [Pizer et al. 1998, Furst and Pizer 1998, Morse et al. 1998, Eberly et al. 1994, Fritsch et al. 1995a] to find medial structures in images. Core tracking methods differ from other skeletonization methods because they operate directly on grayscale images and do not require an explicit specification of the boundary of the object of interest. However, when the boundary is given, core tracking

[4]The A_i^j taxonomy of medial points, due to Giblin and Kimia [2000], was discussed in Sec. 2.1.1



a. b.

FIGURE 2.8. Medial atoms used for core tracking. a. A medial atom defined by position, radius, orientation and object angle. b. An Gaussian derivative filter associated with the atom, as used for core tracking.

can be applied to the characteristic binary image formed by the boundary, in which case the results are analogous to the results of other multi-scale skeletonization methods.

Core tracking methods extract *cores*, which are loci of medial atoms that optimally agree with the image information. Different types of medial atoms have been used in core tracking, including order 0 atoms, order 1 atoms, and even hybrid atoms that include the orientation parameter but fix the object angle to always equal 90 degrees [Stetten and Pizer 1999, Fridman et al. 2003]. For the purpose of this discussion let us restrict ourselves to order 1 medial atoms.

The agreement between a medial atom and an image is measured by a *medialness function*, the domain of which is the set of parameters defining a medial atom. Given a tuple of parameter values, the medialness function measures how well the position and orientation of the boundary nodes of a medial atom defined by these parameters match edge-like structures in the image.

The medialness function most widely used in the core tracking literature is computed using the image intensity gradient at boundary node locations. Given an image I and

an order 1 medial atom $m = \{\mathbf{x}, r, \mathbf{R}, \theta\}$, this function is defined as

$$M(m) = \nabla_\sigma I(\mathbf{y}^{-1}) \cdot \mathbf{U}^{-1} + \nabla_\sigma I(\mathbf{y}^{+1}) \cdot \mathbf{U}^{+1} \ , \tag{8}$$

where $\nabla_\sigma I(\mathbf{y})$ is the image gradient computed at the point $\mathbf{y}$ by convolution with the gradient of the isotropic Gaussian kernel with aperture σ :

$$\nabla_\sigma I(\mathbf{y}) = \int \nabla G_\sigma(\mathbf{y} - \mathbf{z}) I(\mathbf{z}) \ \mathrm{d}\mathbf{z} \ . \tag{9}$$

The aperture σ is proportional to the radius of the medial atom. This proportionality makes the medialness function invariant to the magnification of the structures in the image.

The generic cores in a p-dimensional parameter space are $p-1$ dimensional height ridges of the medialness function. A $p-k$ height ridge of a function f is a continuous locus of points, which are relative maxima of f in $p-k$ linearly independent k-dimensional subspaces of the parameter space. A mountain crest is an example of a $2-1$ height ridge since the crest is a relative maximum in the subspace defined by the direction of the gradient of the altitude function. In general, the vectors defining the $p-k$ subspaces in which f is maximal are themselves dependant on f and are different at distinct points in the parameter space. Core tracking literature has made two different choices for defining these sub-spaces, resulting in two classes of cores: *maximal convexity cores* and *optimal parameter cores*. In maximal convexity cores the subspaces in which the maxima are attained are defined by the $p-1$ directions of greatest second derivative of M, computed at each point as the unit eigenvectors of the Hessian matrix [Eberly 1996]. In optimal parameter cores, the atoms are required to attain a local maximum in radius, orientation, and object angle, as well as in the direction orthogonal to the tangent space of the core in Euclidean space, which is defined by the optimal orientation [Furst and Pizer 1998, Fridman et al. 2003].

Core tracking methods work by following a core from a starting point in the parameter space. A user specifies a location, size, orientation and object angle of an initial medial atom and the algorithm searches for a ridge point in the vicinity of this initialization. The algorithm then tracks the core, taking small steps in the parameter space until some termination condition is met. Core tracking has been implemented using marching cubes methodology [Furst and Pizer 1998, Lorensen and Cline 1987] and using predictor-corrector methods [Fritsch et al. 1995b].

Core tracking methods reported in the literature track a single height ridge and thus yield a medial locus without branches. However, it has been shown that height ridges are connected by other relative critical structures, such as height valleys and saddle manifolds which could be tracked in order to extract a branching medial locus from an image [Damon 1999, Miller 1998, Keller 1999].

Cores, shock loci, and Voronoi skeletons constitute three classes of approaches to approximating the medial locus of an object at any given level of detail. The ability to extract medial loci at a scale makes these three methods robust with respect to small boundary differences. However, these three methods are *deterministic* in the sense that the medial locus is defined entirely by the boundary of the object. These methods do not provide a way of incorporating prior knowledge about the shape of the medial locus into the extraction process. This limitation makes it difficult, if not impossible, to use skeletonization methods to produce features for shape characterization.

2.1.6. Pattern Theory and Deformable Models. Some of the most important driving problems of m-rep research, such as medical image segmentation and shape characterization, fall into the realm of Grenander's *pattern theory* [Mumford 1996, Grenander 1976, 1978, 1981]. Pattern theory provides a mathematical framework for solving a general category of problems that use observations of the world to discover its true state. M-rep methodology uses many of the results of pattern theory.

The fundamental conceptual elements of pattern theory are the world and the observations of it. The world may assume many different states and each state of the world w has some probability $p(w)$ of being the true state. Neither the true state of the world nor the probability distribution on the states are known to the observer.

Observations such as images, signals, or geometric measurements reflect the state of the world in a limited and often obscure fashion. The perceptual process used to produce the observations is error-prone and many different observations can be made from a particular state of the world. The perceptual process is governed by a probability distribution $p(f|w)$, which describes the likelihood of obtaining an observation f given that the world is in state w. In order to compute the likelihood term $p(f|w)$ it is necessary to understand the uncertainty that different stages of the sensory process introduce while forming an observation.

Pattern theory holds that observations in nature are formed and transformed according to a fixed set of rules. The perceived image is the output of a stochastic deformation pipeline at whose input lies the true, uncorrupted image of the world. The possible deformations in this pipeline are classified by pattern theory into the categories of (i) noise and blurring, (ii) superposition, (iii) domain warping, and (iv) interruption. Noise and blurring are caused by the relatively low resolution of our sensors. Superposition is a term that describes the linear or nonlinear combination of several underlying processes to form an observation. The mixing of sounds from different sources or the combination of shape and lighting conditions in producing intensity edges in an image are examples of superposition. Domain warping refers to a diffeomorphism of the image domain caused by the imaging process. Interruption describes interference between observations, such as occlusion of an object by another in an image. Pattern theory goes a far way in describing the mathematical and stochastic properties of these four deformations.

The hypothesis of pattern theory states that in practical problems the loss of information in the sensory process is so great that no algorithm can adequately discover the true

state of the world using feed-forward deterministic analysis of observations. According to pattern theory, an observer must maintain a prior model of the world and adapt it to observations in an iterative analysis-synthesis sequence.

The observer's prior model of the world can be based on past experience of on empirical knowledge. For instance, an anatomist segmenting an organ in a medical image uses her expertise to come up with a general idea of the expected shape of the organ and its likely location in the image. A prior model of the world is expressed as a probability distribution $P(w)$ on the different states of the world w.

Bayesian decision theory provides a framework for updating the observer's prior model of the world $P(w)$ using a new observation f. The updated model of the world is represented is the posterior distribution given by Bayes' rule

$$P(w|f) = \frac{p(f|w)P(w)}{p(f)}. \tag{10}$$

The posterior distribution is the 'best possible guess' at the unknown distribution $p(w)$ that can be made using the observation and the observer's prior knowledge. The term $p(f)$ in the denominator carries little importance: its purpose is to make $P(w|f)$ integrate to 1. The state

$$\hat{w} = \arg\max_w P(w|f) = \arg\max_w \left[p(f|w)P(w)\right].$$

is the optimal guess at the true state of the world and is called the *maximum a posteriori* (MAP) estimate.

Direct computation of the MAP estimate is prohibitively expensive in practical problems where the number of variables needed to describe the state of the world is large. Pattern theory provides a general framework for constructing an algorithm that follows the principles of MAP estimation but does not require explicit computation of the prior and posterior probability distributions. The algorithm must model the following *deformable models framework*:

(1) start with a guess w;

(2) generate synthetic image f_w;

(3) compare f_w to observation f;

(4) update w accordingly;

(5) repeat steps 2 through 5 until convergence.

This generative process resembles human perception: our knowledge of the world is not given at birth, but rather learned by contrasting prior experience with new observations and learning from the differences over and over again.

Let us examine the analysis-synthesis idea in the context of image segmentation. A pattern theoretic algorithm begins with a model of the object of interest. The expected appearance of the object, given the model is then synthesized and compared to the input image. The places in the model where the agreement between the synthesized image and the input image is weak are adjusted and the process is repeated until convergence. The following section described this process in detail in context of m-rep based image segmentation.

A general overview of deformable models is given by McInerney and Terzopoulos [1996]. Deformable model methods differ in the way that they represent objects and in the way that they construct priors. Examples of object representations include point distribution models [Cootes et al. 1995], simplex meshes [Delingette 1994], basis function representations [Staib and Duncan 1992, Székely et al. 1996], and representations based on the interiors of objects [Joshi et al. 1997, Christensen et al. 1997]. Prior terms may be based on statistical learning (e.g., [Cootes et al. 1995]) or on heuristic knowledge about geometric properties of objects, such as smoothness (e.g. [Caselles et al. 1995]).

2.1.7. M-Reps: A Medial Object Representation. The term *m-rep* can refer both to a geometric object representation and to a methodology that uses this representation

for image analysis, following the deformable models framework. This section discusses both aspects of m-reps.

The distinguishing properties of m-reps as an object representation are

- The medial locus of an object is represented explicitly.
- A fuzzy approximate representation of the object's boundary is implied by the medial locus representation.
- An accurate description of the boundary is given by a smooth fine-scale deformation of the fuzzy implied boundary.

The explicit specification of an object's medial locus by m-reps makes it possible to compare similar objects in terms of symmetries. In contrast, medial loci yielded by applying a skeletonization method to similar objects are poorly suited for such comparison because their branching topology is not explicitly specified and may differ drastically, as in the case of nearly circular objects.

M-reps come in different flavors. Discrete m-reps, due to Pizer et al. [1999], represent the medial locus using a structured sparse set of medial atoms. Continuous m-reps, defined in Ch. 5, represent the medial locus as approximation curves or surfaces defined by a set of $\{\mathbf{x}, r\}$ control points. This section describes discrete m-reps.

M-reps specify the figural composition of an object explicitly. An m-rep representing a complex object contains multiple components called *figures*. Each figure is an array of medial atoms. These figures are are organized into a hierarchy of parent-child relationships, with parents representing the substantial parts of the object, such as the palm of the hand, and children representing protrusions and indentations, such as the fingers. The figural graph of an m-rep resembles the composition of a geometric object into figures with joints but does so only at a conceptual level. The manner in which an m-rep is organized into figures is guided by structural and conceptual reasons, rather than by desire to precisely mimic the medial branching topology of the objects being represented.

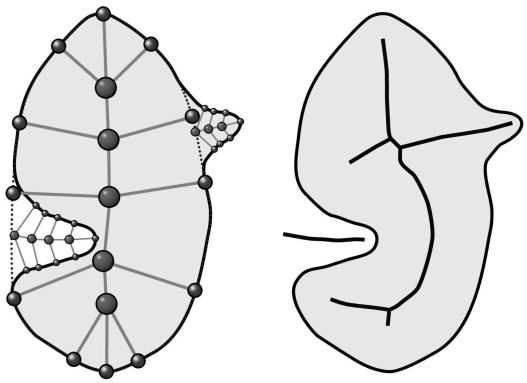
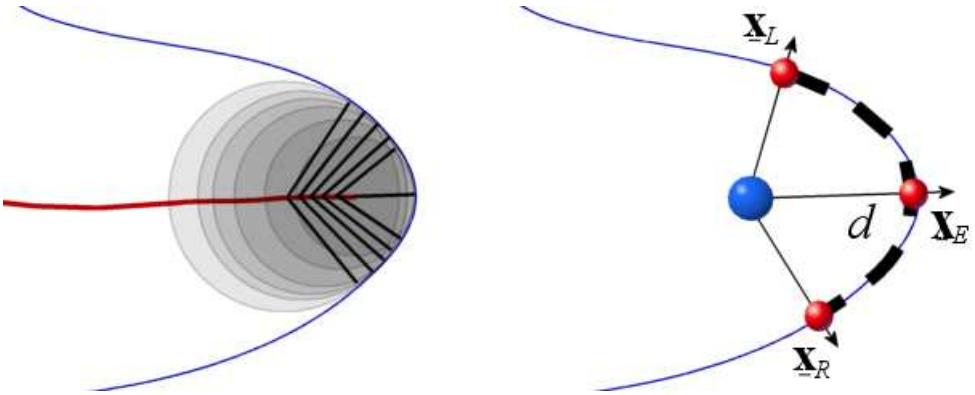


FIGURE 2.9. (Left). Representation of an object using a discrete m-rep. The m-rep is organized into a hierarchy of figures based on structural properties of the object. At the root of the hierarchy lies the main figure whose implied boundary is indicated by the dotted curve. The children of the main figure are the protrusion and indentation figures. (Right). Continuous medial locus of the same object. Branches in the medial locus are determined by the geometry of the boundary and include branches that do not contribute to the structural description of the object.

The difference between the figural composition of m-reps and the branching topology of corresponding objects is illustrated in Fig. 2.9.

An m-rep figure is defined as a lattice of order 1 medial atoms (defined in subsection 2.1.5). In two dimensions medial atoms composing an m-rep figure are organized into a linked list, or chain. In three dimensions atoms are organized into a mesh structure. The edges connecting neighboring atoms in a chain or mesh are called *intra-figural links* in contrast to *inter-figural links* that describe relationships between atoms across different figures.

The locally cylindrical geometry of nearly-tubular three-dimensional structures such as blood vessels is represented by m-rep figures containing chains of special medial atoms. These atoms can be envisioned as a regular medial atom that is spinning around its bisector **b**, or in terms of the rowboat analogy as a kayak doing an Eskimo roll. Aylward and Bullitt [2002] use m-reps consisting of such atoms to model vessel trees.



a. b.

FIGURE 2.10. Medial geometry of end atoms. a. The continuous relationship between a point on a medial curve and the points of contact between a disk inscribed at the point and the boundary of the object asymptotes at the end point of the curve: equal steps along the medial curve result in increasing steps along the boundary. b. To account for this asymptotic relationship m-reps describe ends of figures using special end atoms that describe the entire end-cap of the figure as an arc of an oval.

Another class of special medial atoms is used at ends of chains that form two-dimensional figures and at edges of meshes in three-dimensional figures. These atoms correspond to medial points of type A_3^1, which are located at the ends of medial curves and at edges of medial surfaces and thus have a single-point boundary pre-image. The medial geometry in the neighborhood of an A_3^1 medial point is singular. As we approach such a terminal medial point along the medial curve or along any path on the medial surface, the object angle collapses to zero and its derivative with respect to arc length asymptotes. In terms of the rowboat analogy, the rower approaching an such a point brings the oars together in front of the boat infinitely quickly. This asymptotic behavior is demonstrated in Fig. 2.10a.

The special end atoms used at edges of m-rep figures do not represent the precise A_3^1 medial point, but rather describe a whole section of the object in the neighborhood of this point. In two dimensions an end atom represents the end-cap of a figure using an oval arc, as illustrated in Fig. 2.10b. An end atom is defined as a medial atom that

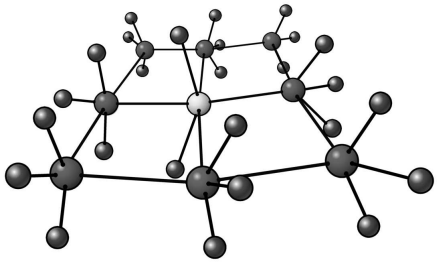


FIGURE 2.11. A diagram of three-dimensional m-rep figure organized as a 3×3 quadrilateral mesh of medial atoms. The atom in the middle of the mesh in a regular medial atom, the rest are end atoms.

contains a third boundary node $\{\mathbf{y}^0, \mathbf{U}^0\}$, which describes the point at the tip of the oval arc. This third boundary node is defined as

$$\begin{aligned} \mathbf{U}_0 &= \mathbf{b} \ , \\ \mathbf{y}_0 &= \mathbf{x} + \eta\, r\, \mathbf{U} \ , \end{aligned} \tag{11}$$

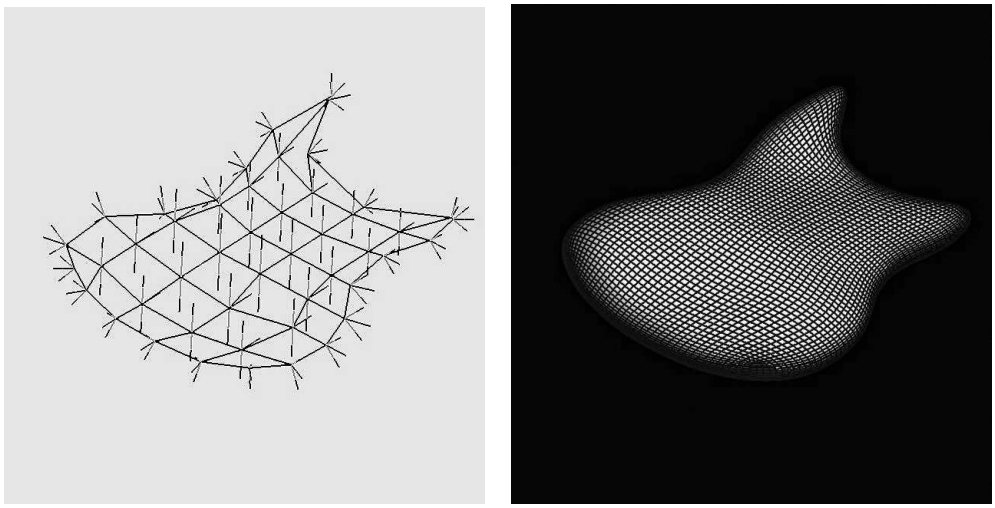
where η is an additional *elongation* parameter that end atoms contain in addition to the standard parameters $\mathbf{x}, r, \mathbf{R}$ and θ. The elongation parameter satisfies $\eta \geq 1$ to ensure that the oval arc described by the end atom is at least as sharp as a circular arc. End atoms in three dimensions appear at edges of meshes of atoms and thus are connected into chains. A chain of end atoms describes the shape of a figure in the neighborhood of a crest of the boundary. Each end atom represents the cross section of this neighborhood in the plane orthogonal to the crest line using an oval described by the elongation parameter η. Fig 2.11a shows diagram of a medial figure in three dimensions as mesh of medial atoms and end atoms.

I stress that the boundary node $\{\mathbf{y}^0, \mathbf{U}^0\}$ is used to describe a tip of a figure using an oval arc or to describe the crest of a figure using a locus of such arcs. The position of the node $\{\mathbf{y}^0, \mathbf{U}^0\}$ need not correspond precisely to the position of the maximum of curvature or the crest on the boundary of the object described by the m-rep. End atoms make it

possible to stabilize the medial object description at the ends of figures. One of the reasons why the skeletonization process is considered unstable is because small changes to the boundary of an object can displace local maxima of boundary curvature by large amounts. Since these maxima correspond to end points of medial curves, the medial locus is sensitive to small boundary changes. By not explicitly specifying the position of the maximum of curvature on the boundary, m-reps can consistently describe similar objects whose exact medial loci differ at the ends of figures. The elongation parameter η allows one to coarsely represent ends of differently shaped figures with more stability than would be possible if the m-rep end-caps were always constrained to be circular (or spherical in 3D), in which case the radii of end atoms would have to match very closely the radii of maximal curvature of the figure and thus, the position of the end atoms would be significantly affected by slight changes to the object's boundary.

By representing the medial locus of an object Ω using a collection of order 1 medial atoms $m_1 \ldots m_N$, an m-rep M explicitly defines a set of corresponding boundary nodes $\{\mathbf{y}_1^\bullet, \mathbf{U}_1^\bullet\} \ldots \{\mathbf{y}_N^\bullet, \mathbf{U}_N^\bullet\}$. By connecting neighboring nodes using continuous patches of curves or surfaces a closed continuous description of $\partial\Omega$ can be constructed. This continuous description is called the *boundary implied by the m-rep M*. The implied boundary can be constructed using Beziér patches or using subdivision surfaces, which, according to Thall [2002], are more appropriate for describing three-dimensional objects. Fig. 2.12b shows the implied boundary of a single-figure three-dimensional m-rep.

In direct contrast to skeletonization methods, m-reps derive the boundary description of an object from its medial description. Since the composition of an m-rep into figures and atoms is explicitly imposed, one can describe different objects using m-reps with the same figural composition. The ability to describe a set of related objects using m-reps with the same figural composition makes it possible to compare objects based on their medial properties. To compare two m-reps that have the same figural composition, one only needs to compare the values of the medial atom parameters, assuming



a. b.

FIGURE 2.12. a. A three-dimensional m-rep figure organized as a quadrilateral mesh of medial atoms. b. The implied boundary constructed by using subdivision surfaces to interpolate the boundary nodes of the medial atoms in the figure. (Courtesy A. Thall.)

that medial atoms represent corresponding locations in the objects. However, to compare medial structures with different branching topologies, such as medial loci produced by skeletonization algorithms, one needs to employ complex tree matching algorithms. Common figural composition makes it possible to estimate probability distributions on the parameters of medial atoms and hence makes medial based shape characterization a reality.

The ability to impose an common figural compositions on m-reps describing a set of similar objects comes at the price of accuracy. The accuracy with which an m-rep describes the boundary of an object in not only limited by the resolution of the m-rep, i.e., the number of atoms used, but also by the explicit a priori specification of the structural composition of the m-rep. For instance, a single figure m-rep that can accurately represent many worm-like objects can not be used to represent an object in the shape of a cross, regardless of the number of atoms used. The work of Styner on modelling various subcortical organs shows that in many practical applications collections

of biological objects can indeed be accurately represented by m-reps sharing a common figural composition [Styner 2001]. However, in applications such as modelling of botanical trees a common figural composition is unfeasible.

To address the issues of accuracy, m-reps methodology uses the idea of *controlled tolerance*. The boundary implied by an m-rep is not considered a fixed curve or surface but rather a maximal set of a probability density defined on the ambient space. This density is called the *fuzzy boundary* of the m-rep.

Earlier m-rep methodology defined the fuzzy boundary by placing one-dimensional Gaussian densities at a grid of points regularly sampled along the implied boundary. Each Gaussian density is oriented in the direction normal to the implied boundary and its standard deviation is proportional to the local width of the object (i.e., the value of r interpolated between the values of the adjacent medial atoms). The proportionality between the radius of a medial atom and the tolerance of the fuzzy boundary at the atom's boundary nodes makes the fuzzy boundary get tighter as the object narrows. The constant of proportionality ρ is specified explicitly for each set of similar objects that must be represented by m-reps with a common figural composition. The coarse-to-fine shape analysis method described in Ch. 3 uses this definition of fuzzy boundary.

The limitation of this early approach Pizer et al. [1999] is its inability to model the correlation in the tolerance of neighboring points on the fuzzy boundary. The model is analogous to white noise, as each Gaussian is uncorrelated with its neighbors in the grid. Neither did the older model take the variability in a population of objects into account when describing boundary tolerance.

In more recent work by Lu et al. [2003b], the fuzzy boundary is described using a probabilistic Markov random field model that is constructed using a training set of objects. In this work, m-reps are treated as a hierarchical representation that describes objects and object complexes at different levels of detail. These levels include the object level, at which the configuration of separate objects in an object complex is described,

the figural level at which relative configuration of the m-rep figures in each object is described, the medial atom level, which describes figures in terms of medial atoms, and the boundary level which describes object boundaries in terms of boundary nodes. The Markov random field model describes the distribution of the residual between primitives used at each level of the representation (e.g. objects, figures, atoms, boundary nodes) and the positions of these primitives implied by the previous level of representation. Following the Markov assumption, the probability distribution of each residual is conditioned only on the residuals of its immediate neighbors. Under this model, the fuzzy boundary of an m-rep is described in terms of the distributions of boundary-to-medial residuals. This fuzzy boundary captures the local variability of boundary nodes in the training set of objects. It also models the correlation between neighboring boundary nodes.

Up to this point I described m-reps as a data structure used to represent objects, i.e, a graph whose vertices are formed by medial atoms and whose edges are formed by inter-figural and intra-figural links. Let us consider the class of operations that can be applied to this data structure. These operations can be categorized into *structural operations* and *structure-invariant operations*. Structural operations add vertices to the graph, remove vertices from it, and change the connectivity between the vertices. Structure-invariant operations only change the values of medial atom parameters at the vertices.

In current m-rep methodology, structural operations on m-reps are performed through user interaction, while structure-invariant operations are used as a step in m-rep based segmentation and other automatic algorithms. In this section I only discuss structure-invariant operations. Among them two special types of operations are of interest: *similarity transformations* and *figural transformations*. Similarity transformations apply translation, magnification, and rotation about a fixed point in space to all the medial atoms in an m-rep. Similarity transformations change the pose of an m-rep without

affecting its geometric form [5]. Figural transformations apply translation, magnification, rotation, and elongation to a single figure in an m-rep. A figural transformation changes the geometric form of the m-rep but, with the exception of elongation, keeps the geometric form of the figures constant.

The algorithm for image segmentation using m-reps follows the deformable models framework defined in the previous subsection. The details of m-rep segmentation methodology can be found in papers by Fritsch et al. [1997], Pizer et al. [1999], and Joshi et al. [2002]. The basic outline of the algorithm for segmenting an object Ω in an image I is the following:

(1) Explicitly specify a template m-rep M with a suitable figural composition. Compute a prior distribution $P(M)$ on the parameters of the m-rep.
(2) Using Bayes rule, compute the match between the template M and the image as the posterior probability density $P(M|I)$
(3) Deform parameters of M using structure-invariant operations in a way that would increase $P(M|I)$.
(4) Repeat steps 2,3 until convergence.

The actual segmentation task is organized into several stages, each following the above four-step procedure. In each stage a different set of structure-invariant operations is used in Step 3. In the earliest stage only similarity transformations are allowed: the template moves, rotates, and changes size relative to the image but retains its geometric form. In the next stage, figural transformations are allowed. Optionally there may be an intermediate stage between the similarity and figural stages in which groups of connected figures are allowed to move relative to each other. This optional stage is used

[5] In this dissertation the term "geometric form" is used instead of the word "shape" to describe the geometric properties of an *individual* object, which do not change when the object is translated, rotated, or uniformly scaled. The term "shape" is used to describe the *collective* geometric form of a population of objects.

when segmenting object complexes, such as groups of organs. In a later stage, individual medial atoms or small groups of them are allowed to deform.

The final state of m-rep segmentation involves a boundary deformation. Let M be an m-rep optimally fitted to an image after the similarity, figural, and atom stages of the deformation. The shape of the implied boundary of M is restricted by the predetermined figural composition of M. In order to better segment the true boundary of the object in the image, the implied boundary is deformed by moving each point on the boundary along the boundary normal, as to maximize a posterior density.

Each stage of the deformation algorithm uses the same image match term $p(I|M)$ (the likelihood term in Bayesian terminology). The image match is computed by integrating a local measure of match over the implied boundary. The local measure can be the strength of the image edge in the direction normal to the boundary, or the correlation in the neighborhood of the boundary point between the image I and some atlas image. In either case, the degree of match is weighted by a Gaussian kernel whose aperture is the same as the local tolerance of the fuzzy boundary: ρr. The Gaussian weighting gives a width-proportional tolerance of the image match measure. This tolerance allows the deforming m-rep to be attracted to the true boundary of the object in the image.

The prior probability density is defined differently in each of the four stages of segmentation. In the similarity stage the prior density uses empirical or statistical knowledge about the position of the object of interest in the image. For instance, the expected position, size, and orientation of the heart in a chest x-ray is known from basic anatomy. In the figural stage, a prior is defined on the relative poses of the figures: when segmenting a hand, we roughly know how the fingers are situated in respect to the palm and to each other. In the atom stage, this density combines the prior knowledge about the geometrical form of each figure with a regularity term that rewards smoothness and orderliness of the medial atom mesh. This smoothness term is based on the empirical

fact that medial loci of objects must follow certain geometric rules. For instance, configurations of atoms whose boundary nodes can not be interpolated by curve or surface patches to form a continuous impled boundary are deemed illegal and assigned a zero prior probability density. During the boundary deformation stage, a regularity prior is applied to the boundary displacements. One of the goals of the statistical shape characterization research reported in this dissertation is to compute prior probabilities for m-rep deformation using a set of training objects.

Much more remains to be said about m-reps and their properties. Later in the dissertation, I will come back to this topic, covering in detail the notions of correspondence, figural coordinate systems, and medial atom interpolation. However, the amount of background on medial methodology that is section covers should prepare the reader for the upcoming chapters and for medial literature in general.

2.2. Shape Characterization

> Ham. Do you see yonder cloud that
> 's almost in shape of a camel?
> Pol. By the mass, and 't is like a
> camel, indeed.
> Ham. Methinks it is like a weasel.
> Pol. It is backed like a weasel.
> Ham. Or like a whale?
> Pol. Very like a whale.
>
> Shakespeare, *Hamlet*

The objective of this section is to introduce the reader to the statistical shape characterization methodology that was motivated in Ch. 1. The section is organized as follows. Subsection 2.2.1 introduces the concept of statistical features and summarizes a number of approaches to extracting features from various object representations. Subsection 2.2.2 describes how statistical methods such as density estimation and principal component analysis are used to model shape variability in a population of objects. Subsection

2.2.3 covers statistical classification of the basis of shape: it describes some common discrimination techniques and discusses their application to clinical diagnosis and disease understanding. Subsection 2.2.4 addresses the problem of describing objects in a population using homologous sets of primitives, known as the problem of correspondence. Finally, subsection 2.2.5 presents feature selection, a technique used to reduce the number of features in high-dimensional low sample size classification problems.

First, a warning regarding notation. In computer vision literature, and in the English language in general, the word "shape" can have multiple distinct meanings. The confusion between the following two is of concern here. "Shape" can refer to the geometric form of a particular instance of an object: i.e, "his nose has an unusual shape." "Shape" can also refer to the geometric properties that make a class of objects distinct, without making a reference to an individual instance: "the nose looks like a tetrahedron with smoothed corners, a pair of holes, and perhaps a bump." For clarity, I always use the word shape in the second sense, sometimes qualifying it by an adjective "collective". I use the term "geometric form" when talking about a particular instance. I would say, "his nose has an unusual geometric form."

In these terms, the goal of shape characterization is to express the shape of a population of objects or to express the differences in shape between two classes of objects. The input to a shape classification method consists of a set of typical instances of the object. This input is called the *training sample*, and it is denoted $\Omega_1 \ldots \Omega_n$. Typically, the objects in the training sample are represented as segmented boundaries, but other object descriptions are possible.

In applications where the population is not separated into classes, the typical objective of shape characterization is to estimate a probability distribution on the space of all geometric forms that can be assumed by an object. When the population is divided into classes, the typical objective is to construct a *classifier* function, which assigns class labels to objects. A good classifier has a high probability of correctly guessing the class

membership of a previously unseen object. One says that such a classifier *generalizes* well.

2.2.1. Representing Objects for Statistical Analysis. In order to make the theories of probability and statistics applicable to the study of object shape, it is necessary to interpret objects in a population as samples from a multivariate random variable. Hence, each object must be represented using a fixed number of homologous measurements. These measurements are called *statistical features*.

The measurements that constitute statistical features must be invariant with respect to translation, rotation, and uniform magnification because these operations do not alter the geometric form of objects. The term *similarity transform* is used to refer to any combination of the above three operations. In other words, statistical features are invariant under the similarity transform.

Shape characterization literature describes many approaches to representing objects using statistical features. In this subsection I will briefly mention some of them, including the landmark representation, the point distribution model, the basis decomposition representation, and the volumetric representation.

The features most commonly used in biological shape analysis are based on *landmarks*. A landmark is "a point of correspondence on each object that matches within and between populations," as defined by Dryden and Mardia [1998], who go on to define three types of landmarks. An *anatomical landmark* is a point of special biological or structural significance. For example, in face analysis, the center of the iris or the corner of the mouth are anatomical landmarks. A *mathematical landmark* is a point with unique geometric properties, such as a singularity or a critical point. A *pseudo-landmark* is a point whose position is derived from the positions of other landmarks, for example a point on the edge of the face located half way between the base of the ear and the tip of the chin. Each type of landmark has its advantages and disadvantages. Anatomical landmarks can be found

consistently on all objects in a population, but their identification is costly because it requires human expertise. Mathematical landmarks are easy to locate automatically, but in many populations critical points and singularities are not found consistently between objects. Pseudo-landmarks are also easy to find but their significance is unclear. For example, a point located on the outline of the hand half way between the tip of the thumb and the tip of the pinky finger may lie on the tip of the middle finger on one person's hand and on the side of a finger on the hand of another person.

In early shape analysis literature statistical features were constructed by taking angles and relative distances between landmarks [Dryden and Mardia 1998, Reyment et al. 1984], and this approach is still widely used. In the work of Kendall [1984], Bookstein [1989], and Dryden and Mardia [1998], features are constructed by taking the coordinates of landmarks after first optimally aligning the objects using the Procrustes algorithm. For each object in the population this algorithm finds a similarity transform that minimizes the mean squared distance between the landmarks on the object and the corresponding landmarks in a template object. Features based on the coordinates of aligned landmarks lie in the Euclidean space and are better suited for statistical analysis than measurements such as angles and relative distances, which are cyclic and non-negative.

The number of anatomical and mathematical landmarks that can be used in most real-world applications is limited by the practical cost of identifying the anatomical landmarks and by the inconsistency of the mathematical landmarks. The cost of producing landmarks is especially prohibitive in three dimensions.

Methods that compute probability distributions on sparse sets of landmarks, such as Bookstein's principal warps method [Bookstein 1989], use interpolation to reconstruct the geometrical form of objects between the landmarks.

In the context of this dissertation, the disadvantage of sparse landmark methods lies in their inherent inability to capture local geometry. Even if some of the landmarks capture truly local information, they do so only at the locations that are *a priori* deemed

to be important. Using sparse landmark methods, it is not possible to find the important local components of geometric variability in a population of objects.

An example of a shape model that uses a dense set of pseudo-landmarks is found in the work of Cootes et al. [1995]. Their *point distribution model*(PDM) is constructed by taking the coordinates of pseudo-landmarks after aligning them using the Procrustes algorithm. In early research, these landmarks were obtained by manually identifying a few anatomical landmarks on the boundary of each object and by densely sampling points located at equal intervals along the boundary between the anatomical landmarks. Fully automatic pseudo-landmark selection requires an algorithm that can identify corresponding locations across a population of objects. The development of such an algorithm is perhaps the most important and difficult problem facing the shape characterization field. Subsection 2.2.4 covers this problem and the recent advances toward its solution.

Cootes et al. use PDM as a part of a combined model that describes both the shape of objects and their appearance in images. In their work on Active Shape Models [Cootes et al. 1995], appearance is represented by a statistical description of image intensity in the neighborhood of each landmark in the PDM. In more recent work on Active Appearance Models, Cootes et al. [1998] compute a statistical model of image intensity over an set of image voxels that are enclosed by the boundary of an object or lie in a band surrounding the boundary.

Statistical features based on points densely sampled from interiors of objects are also used in shape characterization methods developed by Davatzikos et al. [1995], Christensen et al. [1997], Joshi et al. [1997], Csernansky et al. [1998], and others. These methods derive statistical features from deformation fields that optimally warp a template object to each object in the population. The warp is based on maximizing the image intensity match between the warped template and each object, while adhering to physical constraints, such as rules of elastic deformations or rules of fluid dynamics. Displacement vectors sampled from deformation fields serve as statistical features for shape characterization.

The above methods densely sample the boundary or volumetric object representation to produce local features. A different class of methods computes features by decomposing the boundary of each object into a set of basis functions. For instance, Staib and Duncan [1992] decompose boundaries of three-dimensional objects into Fourier surfaces and use the first few Fourier coefficients as features. Similarly, Székely et al. [1996] use spherical harmonics as the basis into which they decompose three-dimensional object boundaries. Unlike the statistical feature derived from landmarks, features computed as coefficients of basis functions have global support: a local change to the geometrical form of an object can change the values of all of the features describing the object. However, features with local support can be derived by regularly sampling Fourier surfaces.

In this subsection I gave examples of how statistical features can be computed from object representations. In the following subsection I will discuss how these features are used to model shape.

2.2.2. Shape Density Estimation and PCA. Let $\Omega_1 \ldots \Omega_n$ be a training sample drawn from a population of objects that are described using k features. Let each object Ω_i be described by the vector of values

$$\mathbf{a}_i = [a_{i1}, \ldots, a_{ik}]. \tag{12}$$

In the case of landmark-based object representations, these values are formed by the grouped coordinates of the aligned landmarks: $a_{i1} = x_{i1}, a_{i2} = y_{i1}, a_{i3} = x_{i2}, a_{i4} = y_{i2}, \ldots$. The matrix $A = \{a_{ij}\}$ whose n rows are the transposed feature vectors is called the *feature matrix*.

Shape characterization problems treat the feature vectors $\mathbf{a}_1 \ldots \mathbf{a}_n$ as a sample from a random variable $\mathbf{a}$. Standard statistical techniques can estimate the probability density of $\mathbf{a}$ using the sample. A common approach is to model $\mathbf{a}$ with the multivariate normal

distribution whose parameters are estimated from the training sample:

$$\mathbf{a} \sim N(\bar{\mathbf{a}}, \mathbf{S}), \tag{13}$$

where $\bar{\mathbf{a}}$ is the sample mean, and $\mathbf{S}$ is the sample covariance matrix.

Under the Gaussian assumption, the sample mean represents the typical or average geometric form of the objects in the training sample. However, when the features have a non-Gaussian distribution, the geometric form corresponding to the sample mean may be that of a very unlikely object, or it may not be a valid geometric form at all. For instance, if the underlying distribution forms a two-dimensional ring in feature space, the mean is the center of the ring and is not a representative of the training sample.

The components of the random variable $\mathbf{a}$ can be strongly correlated, especially in shape characterization problems that use densely sampled features. For example, the coordinates of a pseudo-landmark sampled along the outline of a smooth object strongly correlate to the coordinates of its neighbors. In such cases the variability in the training sample can be accurately captured using only a few appropriately chosen uncorrelated univariate random variables. These variables can be found using a statistical technique called *principal components analysis* (PCA).

PCA expresses the random variable $\mathbf{a}$ as a weighted sum of $n-1$ uncorrelated univariate random variables $\beta_1 \ldots \beta_{n-1}$, which have zero means and unit variances:

$$\mathbf{a} = \bar{\mathbf{a}} + \sum_{j=1}^{n-1} \beta_j \mathbf{v}_j \ . \tag{14}$$

The vectors $\mathbf{v}_1, \ldots \mathbf{v}_{n-1}$ define orthogonal directions in feature space space that are called *principal components*.

Given a random sample $\mathbf{a}_1, \ldots, \mathbf{a}_n$, the principal components are computed in terms of eigenvalues and eigenvectors of the sample covariance matrix. The vector $\mathbf{v}_i$ is given

by multiplying the square root of the i-th eigenvalue (where eigenvalues are arranged in order from largest to smallest) by the corresponding i-th eigenvector.

For any $d < n$, the subspace of the feature space that passes through the mean $\bar{\mathbf{a}}$ and is spanned by the first d principal components is the best d-dimensional mean square fit to $\mathbf{a}_1, \ldots, \mathbf{a}_n$. For example, $\mathbf{v}_1$ coincides with the line that passes through $\bar{\mathbf{a}}$ and has the smallest mean square distance to the training vectors, $\mathbf{v}_2$ has the same property in the subspace orthogonal to $\mathbf{v}_1$, and so on.

Typically, when PCA is applied to correlated features, only the first few principal components have lengths that are an order of magnitude above the standard deviation of the noise that is inherently present in the feature extraction process. Said less formally, only a few components are needed to express the variability in the 'signal', and the residual variability that is not captured by those few components may as well be 'noise'.

In their active shape models algorithm, Cootes et al. [1995] apply PCA to features computed from a dense pseudo-landmark object representation. Landmark representations make it possible to reconstruct the geometric form associated with each point in feature space. By reconstructing a sequence of geometric forms corresponding to points sampled along one of the first principal components, Cootes et al. produce visual descriptions of the principal modes of shape variability. Different principal modes may correspond to geometrically or biologically distinct deformations: for example, one mode may account for bending and another for thickening. However, more often than not, the geometric deformation associated with a principal mode consists of a mix of different geometric operations, and it need not be indicative of some underlying biological process. PCA-based visualization of shape variability can be used to construct medical atlases that show typical variations in the shape of human anatomy. [Thompson and Toga 2002, Christensen et al. 1996].

In shape characterization problems where the correlation between features is high, PCA is used to reduce the dimensionality of the problem. As a preprocessing step, PCA

can be used to take a thousand-dimensional problem and make it a ten-dimensional one. In machine learning terminology, PCA is called a *feature recombination* technique. Feature recombination techniques like PCA have a disadvantage: new features yielded by them are weighted sums of all of the original features, so each one of them is a global shape measurement. Thus, PCA loses the locality of the original features.

PCA is most effective when applied to features that are normally distributed. In that case, the uncorrelated random variables $\beta_1 \ldots \beta_{n-1}$ are indeed independent. The principal components describe independent components of shape variability and can be used to find the biological processes responsible for this variability. However, when the features are not normally distributed, the geometrical forms corresponding to the traversal of the principal components may be very unlikely to occur in nature, or they may be geometrically invalid. For instance, if the feature vectors form two equal size clusters in feature space, the geometrical form corresponding to the mean of the feature vectors may not resemble the objects in either of the clusters, and it is by no means the 'most likely geometrical form'.

The normal distribution is commonly used to model shape because it frequently occurs in nature, because it has nice mathematical properties, and because it requires a relatively small number of parameters to define. More elaborate models of the feature distributions are possible. For example, Aylward [1997] uses cores, which are similar to those used for image analysis, in feature space to construct continuous mixture models. Such models use more parameters than the simpler Gaussian model, and in order to generalize well, they require a large training sample.

2.2.3. Shape-Based Classification. This section discusses the problem of shape-based classification, which deals with using the geometric form of objects as a basis for automatically placing objects into categories. The input to a classification problem consists of a training set of tuples $\{\mathbf{a}_1, y_1\} \ldots \{\mathbf{a}_n, y_n\}$. Each tuple consists of a feature vector $\mathbf{a}_i$

that describes an object Ω_i and of an associated label y_i that indicates the class to which Ω_i belongs. The goal of shape-based classification is to compute a class membership function c, which, given a previously unknown object, returns the correct class label.

The quality of a classifier can be tested using a second set of objects with known class membership. Let this training set be denoted as tuples $\{\mathbf{a}'_1, y'_1\} \ldots \{\mathbf{a}'_{n'}, y'_{n'}\}$. The error rate of the classifier c is computed as

$$\epsilon = \frac{1}{n'} \sum_{i=1}^{n'} L\left(y'_i, c(\Omega'_i)\right) \tag{15}$$

where $L(y_i, y_j)$ is the loss function, which describes the cost associated with incorrectly misclassifying an object of class y_i as belonging to class y_j. The simplest loss function, called the 0-1 *loss*, assigns the cost of 1 to a misclassification and a cost of 0 to a correct classification. Other loss functions become necessary in diagnostic problems where the cost associated with incorrectly classifying a diseased person as healthy is much greater than the cost of labelling a healthy person as diseased. A classifier is said to *generalize* well if it achieves a low error rate when applied to test data. A classifier that achieves a low error rate when tested using the training data but generalizes poorly to other test data is said to be *overfitted*.

A classifier divides the feature space into regions over which the classifier attains constant values. The boundaries between these regions are called *decision boundaries*. Points on decision boundaries are as likely to belong to one class as to another. Examples of decision boundaries of different classifiers are shown in Fig. 2.13.

The description of a large number of classification methods and a discussion of many related issues can be found in Duda et al. [2001]. One common classification approach involves using the training sample to estimate class-conditional probability densities $p(\mathbf{a}|c(\mathbf{a}) = 1) \ldots p(\mathbf{a}|c(\mathbf{a}) = N_c)$. The classifier c is then computed as the following a

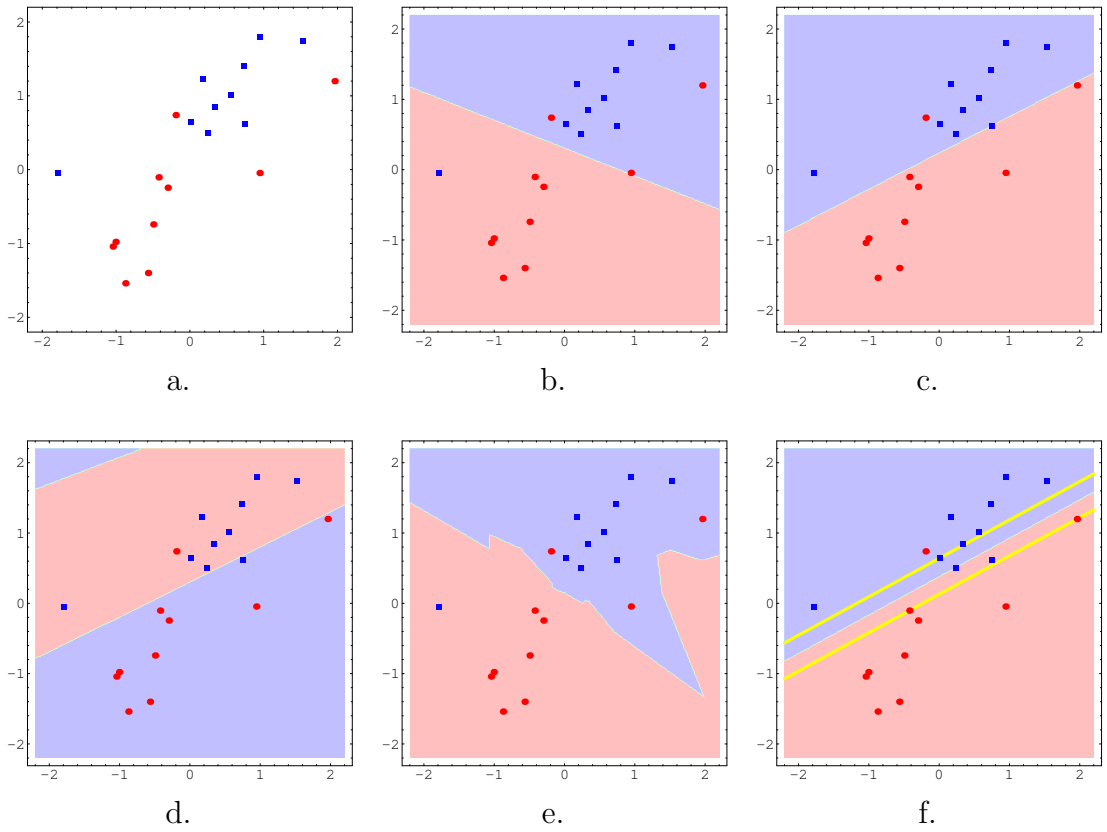


FIGURE 2.13. Two-dimensional classification example. a. A training set consisting of random samples of size 10 taken from two bivariate normal distributions, indicated by red dots and blue squares. b. Difference of means classifier constructed on the training set. The pink and light blue colors indicate the class label assigned by the classifier to each point in feature space. c. Linear classifier based on the Fisher linear discriminant. d. Maximum likelihood classifier for Gaussian distributions with non-equal covariances. e. k-nearest neighbor classifier with $k = 3$. f. Support vector machine classifier with margin denoted by yellow bold lines.

posteriori likelihood estimate:

$$c(\mathbf{a}) = \arg\max_i p(\mathbf{a}|c(\mathbf{a}) = i)P(c(\mathbf{a}) = i) \ , \tag{16}$$

where the probabilities $P(c(\mathbf{a}) = i)$ are based on prior knowledge of the frequency with which the classes are expected to occur. In the absence of this information the prior

probabilities $P(c(\mathbf{a}) = i)$ are taken to be equal and c is called a maximum likelihood classifier.

The Fisher linear discriminant is an example of a classifier that is commonly used in shape classification literature. It is the maximum likelihood classifier for a two-class problem when the class-conditional densities are modelled with a pair of multivariate Gaussian distributions that share a common covariance matrix. The decision boundary of the Fisher linear discriminant is a hyperplane that separates estimated means of the two class-conditional densities. The discriminant is optimal when the true class conditional densities are indeed Gaussian and identically covariant, but when they are not, this classifier may generalize poorly.

Accurate estimation of class-conditional densities requires the size of the training sample to be large with respect to the number of number of features that describe each object. Classifiers that rely on these densities tend to generalize poorly in the high-dimensional low sample size (HDLSS) situation. Classifiers that use Gaussian density models tend to be sensitive to outliers, i.e., elements of the training sample that have a very low probability of occurrence or have been corrupted by the feature extraction process. Non-parametric classifiers tend to overfit the training data.

The *Support vector machines* (SVM) method developed by Vapnik [Vapnik 1995, 1998] is a classification approach that has become a popular choice in shape classification. SVM methodology does not estimate class-conditional densities. Instead, it looks for a pair of parallel hyperplanes, called a *margin*, that optimally separates the elements in a two-class training sample into two groups with different class labels. If there exists a margin that can perfectly separate the training sample, the widest possible margin that does so is considered optimal. If such a margin does not exist, i.e., the groups are inseparable, the optimality is based on the combination of rewarding the width of the margin and penalizing the number of wrongly separated elements. The optimal margin passes through some of the elements of the training sample; these elements are called

support vectors. Support vectors are the elements of the training sample that are most difficult to separate and thus, according to the SVM point of view, should contribute the most to constructing the classifier. The decision boundary of an SVM classifier is formed by the hyperplane that passes in the middle between the hyperplanes forming the margin, as illustrated in Fig. 2.13f.

Generally SVMs are less affected by outliers than the methods that estimate conditional densities. They generalize well in many HDLSS problems where other methods fail. However, the number of support vectors is generically equal to the number of features plus one, and in cases where the number of samples is roughly the same as the number of features, almost every element of the training set becomes a support vector. In these cases SVMs overfit the training data as do the other approaches. The HDLSS situation presents a challenge to all kinds of classifiers. Subsection 2.2.5 addresses this issue by discussing a number of methods that can reduce the dimensionality of the training data prior to classifying it.

Shape-based classification has been used by clinical researchers to relate shape differences with disease progress. For instance, researchers have confirmed that schizophrenia and Alzheimer's disease affect the shape of the hippocampus. Csernansky et al. [1998] obtain this result using features that characterize deformations of a template object into other objects in the population. Shenton et al. [2002] obtain a similar finding for schizophrenia and the amygdala-hippocampus complex using features derived from the spherical harmonic object representation. Styner et al. [2003] confirms these results with features derived from a hybrid object representation that uses m-reps to describe objects at a coarse scale and uses spherical harmonics to describe them at a fine scale. Such results may eventually lead to shape-based classification being used for early diagnosis of schizophrenia. Gerig et al. [2001] and Styner [2001] study subcortical structures of

discordant twins (one twin is afflicted by a disease, and the other is healthy) and demonstrate that differences in shape are better detectable and more correlated with disease than are volume differences.

In theory, shape-based classification can also be used to discover biological processes that are responsible for making the shape of an organ affected by a disease different from the shape of a healthy organ. Shenton et al. [2002] write

> For example, there is evidence to suggest that shape deformations may be associated with the physical properties of morphogenetic mechanisms that directly impact on the particular shape of brain regions during neurodevelopment (Van Essen [1997]; Van Essen and Drury [1997]; Van Essen et al. [1998]). The physical tension of brain growth during neurodevelopment may lead to shape deformations that might be observed using shape measures of brain structures.

A shape classification algorithm that detects a difference between classes of objects and then describes this difference in terms of simple geometrical processes such as thickening, bending, or bump formation can reveal the location and scope of these "morphogenetic mechanisms". For example, if a classification algorithm shows that some disease causes some organ to develop a bump at some particular location, then clinicians can focus their attention on the biological processes that cause tissue growth at that location in that organ.

Subsequently, there recently has been a push in developing methods that qualitatively describe shape differences between classes of biological objects. For example, the algorithm developed by Golland et al. [2001, 2002] uses the gradient of a classifier function taken with respect to the individual features to compute a deformation, which corresponds to the differences between a pair of classes but ignores intra-class variability.

Visual examination of the deformation reveals areas of the object where differences between classes are most pronounced.

2.2.4. Issues of Correspondence. In the context of landmark descriptions, correspondence is the problem of automatically finding a set of landmarks that are consistent across a population of objects. In a more general context, it is a problem of extracting a set of corresponding measurements from objects in a population. The problem of correspondence can be very difficult or impossible to solve, especially when geometry is the only kind of information available. For example, the correspondence between hand outlines can be established fairly well as long as every hand has four fingers and a thumb: mathematical landmarks can be automatically selected at tips and bases of fingers, and between them pseudo-landmarks can be placed at regular intervals. Such an approach would fail miserably if we used it to extract landmarks on the surface of the cerebral cortex because very few mathematical landmarks can be found consistently across a population of cortical surfaces. Points sampled between such mathematical landmarks on the surface of the cortex would fall on the gyri in some instances and sulci in others, and they could by no means be considered corresponding.

The problem of correspondence can be inherently discrete, or it can be continuous. In the discrete case, each object in the population is represented using a finite set of primitives that can not be moved or resampled. Methodology such as the SoftAssign algorithm by Rangarajan et al. [1997] has been developed for combinatorially matching up corresponding primitives, possibly ignoring some primitives that do not match up well.

The continuous correspondence problem occurs when objects in a population are described with a *continuous parameterized representation*, i.e., a continuous mapping from a parameter space to the object space. Examples of continuous parameterized representations include Fourier and spherical harmonics shape descriptions [Staib and Duncan

1992, Kelemen et al. 1999], splines, and continuous m-reps described in this dissertation. In statistical analysis, a set of points in the parameter space can form the features and the problem of correspondence can be recast as the following reparameterization problem.

Let $\mathcal{R}_1(\mathbf{p}), \ldots, \mathcal{R}_n(\mathbf{p})$ be set of parameterized continuous representations of objects $\Omega_1, \ldots, \Omega_n$. The representation $\mathcal{R}_i(\mathbf{p})$ can be a mapping from a domain in $\mathbb{R}^2$ to the continuous surface of a three-dimensional object Ω_i, a mapping from a domain in $\mathbb{R}^3$ to the interior of Ω_i, or a mapping from a domain in $\mathbb{R}^2 \times \mathbb{Z}^+$ to the object's continuous medial locus. The problem of establishing correspondence between the objects is equivalent to finding a set of diffeomorphic mappings $\phi_1, \ldots, \phi_n$ that maximize over all values of the parameter $\mathbf{p}$ some local measure of match between the n representations:

$$\phi_1, \ldots, \phi_n = \arg \max_{\phi_1, \ldots, \phi_n} \int M\left[\mathcal{R}_1(\phi_1(\mathbf{p})), \ldots, \mathcal{R}_n(\phi_n(\mathbf{p}))\right] \, \mathrm{d}\mathbf{p} \; . \tag{17}$$

A simple and commonly used choice is to base the mappings $\phi_1, \ldots, \phi_n$ on the arclength parameterization of the surfaces or curves $\mathcal{R}_1, \ldots, \mathcal{R}_n$. The arclength parameterization is uniquely defined, up to the choice of a point of origin and the directions of the level curves of at the origin. Arclength parameterization is used to establish correspondence for the SPHARM object representation, which is based on spherical harmonics [Brechbühler et al. 1995]. Arclength parameterization provides a heuristic solution to the problem of correspondence, but it makes no guarantee that the geometric or anatomic properties of the points deemed corresponding are indeed similar. Arclength parameterization of the cortex makes no effort to match sulci to sulci and gyri to gyri. However, Davies has reported that that the arclength parameterization used to define spherical harmonics representation of a population of hippocampi yields a correspondence that comes very close to his definition of optimal correspondence Davies et al. [2002], which is described later in this section.

The work of Tagare et al. [1995], Tagare [1997] is an example of a correspondence approach that follows the form of the Equation 17. His work applies to two-dimensional objects whose boundaries are represented by closed curves. The function M, which measures the match between points $\mathcal{R}_1(\phi_1(\mathbf{p})), \ldots, \mathcal{R}_n(\phi_n(\mathbf{p})$ is computed as a function of similarity between discrete derivatives of normal vectors at these points. Another example can be found in the work of Bello and Colchester [1998], who treat the continuous correspondence problem as a problem of non-rigid registration and use the mutual information metric as a measure of correspondence. Methods that derive statistical features deformation from deformation fields [Davatzikos et al. 1995, Christensen et al. 1997, Joshi et al. 1997, Csernansky et al. 1998] implicitly solve the correspondence problem in the form given by Equation 17 because the deformation fields themselves are the optimal mappings $\phi_1, \ldots, \phi_n$.

A different formulation of the correspondence problem can be found in the PDM literature. Kotcheff and Taylor [1997] consider the correspondence optimal if it yields a set of features whose variability in the first few principal components is maximal. Davies et al. [2002] treat correspondence as a minimum description length problem. They search for reparameterizations $\phi_1, \ldots, \phi_n$ that allow the PDM to be expressed using the shortest possible message. The length of this message directly relates to having as much variability as possible occur in the first few principal components while keeping the variability in the remaining components on the order of imaging noise.

2.2.5. Feature Selection. In classification problems, the generalization ability of a classifier can be improved by removing *irrelevant* features, i.e., features that do not contribute to classification. A number of algorithms for automatic feature selection have been developed in the machine learning literature.

Feature selection methods can be categorized into *filter* methods and *wrapper* methods. Filter methods deal with the feature selection task independently of the classification: they find and remove irrelevant features first, and pass the rest on to the classification [Kira and Rendell 1992]. Wrapper methods use classification as a sub-task; as they try different subsets of features, they perform classification and cross-validation on each subset, until an optimum is found [John et al. 1994, Kohavi and John 1997]. Wrapper methods generally perform better than filter methods but are much more time consuming, as each iteration of the method requires an execution and testing of the classification method.

Feature selection methods can also be categorized as *exhaustive*, *randomized* or *sequential*, based on the search algorithm that they employ for finding the optimal feature subset [Aha and Bankert 1995, Jain et al. 2000]. Exhaustive methods search for an optimal subset of n features using either a combinatorial search of all the 2^n possible subsets, or using AI techniques, such as the branch and bound algorithm [Narendra and Fukunaga 1977]. Sequential feature selection methods achieve polynomial time complexity by iteratively adding and subtracting features to a subset in a greedy fashion. Randomized methods employ stochastic search techniques, such as simulated annealing and genetic algorithms. Comparisons of a number of popular feature selection techniques can be found in [Aha and Bankert 1995], and [Jain and Zongker 1997]. According to [Cover and Campenhout 1977, Jain et al. 2000], only the exhaustive search procedure can be guaranteed to produce the globally optimal feature subset.

CHAPTER 3

A Coarse-to-Fine Approach to Shape Characterization in 2D

> Beauty depends on size as well as symmetry. No very small animal can be beautiful, for looking at it takes so small a portion of time that the impression of it will be confused. Nor can any very large one, for a whole view of it cannot be had at once, and so there will be no unity and completeness.
>
> Aristotle, *Poetics*

The work described in this chapter was done in collaboration with Dr. Stephen M. Pizer, Dr. Sarang Joshi, and Dr. J. S. Marron at the University of North Carolina who contributed to the conception and implementation of the research. The experimental results reported below were made possible thanks to the contribution of clinical data by Dr. Guido Gerig, who provided the corpus callosum data from a study of schizophrenia.

3.1. Introduction

This chapter presents a novel coarse-to-fine approach to solving the problem of shape characterization, which is described in detail in Sec. 2.2. Given a population of similar objects, which are represented as images, the problem requires the use of statistical techniques to analyze the principal components of variability in this population or to find differences between distinct subgroups of the population. For example, given images of the corpus callosum of patients in a schizophrenia study, the problem may be to analyze

the typicality and variability of the corpus callosum shape and describe differences in this shape between schizophrenics and healthy subjects.

The coarse-to-fine approach is based on the assumption that form perception and understanding are multilevel hierarchical processes. Often, the overall geometrical form of an object can be perceived when the fine details of the object's boundary are undetectable or distorted. For example, we can tell a potato when we see it from ten feet away. On the other hand, careful examination of the fine details of the boundary may be necessary to distinguish between two similar objects. Moreover, the examination of fine-scale geometric properties of objects can often be performed independently at different object regions. Thus, even though we can see a potato from far away, we need to examine it more closely to tell whether or not it is rotten. The close examination is performed by focusing in turn on each part of the potato's surface, not by simultaneous analysis of the surface in its entirety. Traditional shape characterization methods, such as active shape models, indeed would examine the whole potato at once, using a fine scale description of its surface.

The coarse-to-fine approach uses m-reps to describe the geometric form of objects at different scales. It then analyzes the variability in the global shape of objects using a coarse-scale m-rep description, and it uses the fine-scale description to separately analyze the local shape variability at different regions of objects. By separating the shape characterization problem into a global subproblem and a set of local subproblems, the new approach reduces the number of features used in each subproblem, making shape characterization more manageable in the high-dimensional low sample size setting.

Statistical features used in each coarse-to-fine shape characterization sub-problem are based on relations between neighboring medial atoms. Recall from Sec. 2.1.5 that a medial atom is formed by components that approximately describe curves (or surfaces in 3D) that form the medial loci of objects, as well as by components that approximately describe the radius scalar field defined over the medial locus. By basing statistical features

on these components, it is possible to express variability in the shape of an object in terms of bending, widening, and elongation of the m-rep figures composing the object. Such terms are often more intuitive than the purely local geometrical properties, such as curvature, which can be inferred from boundary representations. Features used to measure local shape variability are expressed in terms of the residual difference between the fine-scale description of an object region and the coarse-scale description of the same region. Such residual features capture just the new information that is introduced by going from a coarse-scale description to the fine-scale description.

When applied to the problem of shape-based discrimination, the new approach creates a hierarchy of coarse-scale global and fine-scale local classification subproblems. By measuring and comparing how significantly the classes are separated in each one of these subproblems, it is possible to gain a qualitative understanding of the morphological process or processes that cause the classes to differ in shape. For instance, if discrimination between the coarse-scale descriptions of the classes is more significant than the localized fine-scale discriminations, we may conclude that the morphological process acts globally on the objects. On the other hand, if the fine-scale discrimination associated with one of the object regions proves to be most significant, that would in effect localize the morphological process to that object region.

In this chapter, the coarse-to-fine approach is applied to shape characterization of the corpus callosum in two dimensions. Results are reported for both a clinical data set and a synthetic data set. The clinical data set is used for visualizing components of shape variability. It consists of binary segmentations of the corpus callosum in the midsagittal MRI projections of 71 subjects in a schizophrenia study. The synthetic data set is used for the purpose of studying discrimination, and it is based on the corpus callosum images in the clinical data set. The corpus callosum images in the synthetic data set fall into three classes. Characteristic representatives of each of the classes are shown in Fig. 3.1. The coarse-to-fine approach can detect that the three classes are different, and it can

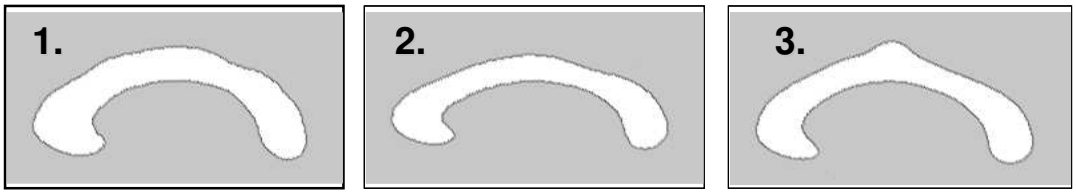


FIGURE 3.1. Representatives of three classes of corpus callosum shapes. Class 1 differs from class 2 in terms of global geometry, and the difference can be detected by examining the classes at a coarse level of detail. The difference between class 1 and class 3 is exhibited only at the middle of the object, i.e., is local. An examination at a finer level of detail may be needed to detect such local differences.

show that the difference between classes 1 and 2 is global, while the difference between classes 1 and 3 is exhibited locally at the corpus callosum midbody (the middle section of the structure).

The methodology presented in this chapter only deals with simple two-dimensional objects like the corpus callosum, which can be described accurately using an m-rep consisting of a single medial figure (i.e., one chain of medial atoms). However, many of the ideas presented here straightforwardly extend to multi-figure objects and to three dimensions.

The following sections describe the details of the coarse-to-fine approach and the experimental results. Section 3.2 shows how a coarse-to-fine family of m-reps is fitted to images of objects in a population. Section 3.3 discusses the construction of statistical features on the basis of m-reps and shows the experimental results of shape characterization using the new approach. Finally, Sec. 3.4 gives a brief discussion of the many issues brought up by this chapter.

3.2. Coarse-To-Fine Object Representation using Discrete M-Reps

The coarse-to-fine approach to shape characterization describes objects in a population using sets of measurements taken at different scales (i.e., levels of detail). Each

object is described with a set of coarse-scale measurements, a set of fine-scale measurements, and possibly, a range of intermediate sets of measurements. M-rep methodology allows such sets of measurements to be computed in a consistent manner from images of objects because, as mentioned in Sec. 2.1.7, the level of detail at which m-reps describe objects can be adjusted explicitly. The following paragraphs describe how this adjustment is made in the process of deformable template-based image segmentation.

Let $\mathbf{M}$ be a single-figure m-rep consisting of medial atoms $\mathbf{m}_1 \ldots \mathbf{m}_n$. A discrete implied boundary of M is formed by the sequence of boundary nodes of the medial atoms

$$\mathbf{x}_1^0, \mathbf{x}_1^1, \ldots, \mathbf{x}_n^1, \mathbf{x}_n^0, \mathbf{x}_n^{-1}, \ldots, \mathbf{x}_1^{-1} \quad . \tag{18}$$

Recall from Sec 2.1.5 that each atom has a pair of boundary nodes with superscripts ± 1, which correspond to the tips of the oars, and the first and the last atoms in a figure have an additional boundary node with superscript 0, which corresponds to the tip of the object. A continuous representation of the implied boundary of $\mathbf{M}$ can be constructed by interpolating this sequence of boundary nodes using Hermite splines.

As described in Sec. 2.1.7, m-rep segmentation is implemented by fitting a template m-rep to an image in a Bayesian energy minimization framework. The energy consists of a prior term that measures the typicality of the deforming m-rep and a likelihood term that measures the match between the boundary of the deforming m-rep and the edge structures in the image. This image match term is implemented by sampling the continuous implied boundary of the deforming m-rep at many points and by measuring the agreement between image intensities in the neighborhoods of these points and Gaussian derivative kernels oriented orthogonally to the implied boundary.

Let $\mathbf{x}_i$ be a point sampled from the continuous implied boundary of $\mathbf{M}$ and let $\mathbf{n}_i$ be the unit normal to the implied boundary at that point. The measure of match between

$\mathbf{M}$ and the image at x_i is computed as

$$[\mathbf{I} * \nabla G_\sigma](\mathbf{x}_i) \cdot \mathbf{n}_i \quad , \tag{19}$$

where $\mathbf{I}$ denotes the image and $*$ is the convolution operator. The standard deviation σ of the Gaussian derivative kernel is taken to be proportional to the local width r_i of the m-rep near x_i:

$$\sigma_i = \rho r_i. \tag{20}$$

This width is computed by interpolating the r values of the medial atoms whose boundary nodes form the segment of the continuous implied boundary on which x_i lies. Using distances that are object-width proportional is in keeping with the basic requirement that shape at any level of scale must be magnification invariant.

By applying m-rep segmentation to the same image using different values of the constant of proportionality ρ, one can produce a family of m-reps that describe the object in the image at different levels of detail. As ρ decreases, the m-reps become more sensitive to local boundary details. An m-rep produced for a given value of ρ can serve as the initialization to the deformable segmentation that uses the next smaller value of ρ.

Following basic intuition from sampling theory, the number of medial atoms used to represent objects at finer levels of detail should be larger than the number of atoms used at the coarser level. Indeed, Gaussian smoothing performed during the segmentation process causes the more high-frequency shape information to be lost for larger values of ρ, thus requiring fewer samples according to Nyquist law.

However, in order to compute residuals between m-reps fitted to an object at different levels of scale, it is necessary to have for each atom on the fine m-rep a corresponding "point of reference" atom on the coarser m-rep. This reference atom is computed by

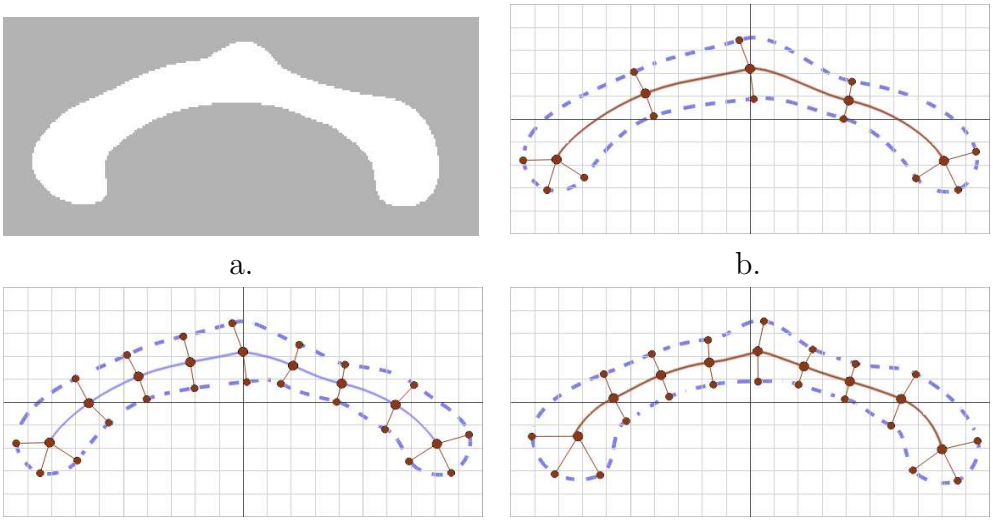


FIGURE 3.2. Examples of coarse and fine m-reps. a. A typical simulated corpus callosum image. b. A coarse m-rep M_0 fitted to the image. c. A prediction m-rep M_0' formed by interpolating the coarse m-rep. d. A fine m-rep M_1 fitted to the image at a higher level of detail.

interpolating the atoms forming the coarse m-rep using a technique described in Sec. 3.2.1. For example, Fig. 3.2b shows a coarse-scale m-rep that describes the corpus callosum using 5 medial atoms and Fig. 3.2c shows a fine-scale m-rep that describes it using 9 atoms. In order to compute the residuals between the two m-reps, the atoms in the coarse m-rep are interpolated forming a continuous locus, and the locus is resampled yielding an m-rep with 9 medial atoms, shown in Fig. 3.2d. The residuals are computed by comparing the 9 interpolated medial atoms to the 9 medial atoms that form the fine m-rep.

The following is the summary of the procedure that is used to compute a coarse-to-fine family of m-reps describing an object in an image. A template m-rep with a few medial atoms is constructed manually. This template is fitted to the image using a large value of ρ, resulting in a coarse level m-rep denoted M_0. Next, the atoms in M_0 are interpolated and the interpolation is resampled to a higher resolution, forming an m-rep denoted M_0', which is called the *prediction m-rep*. M_0' is again fitted to the image, this time using a

smaller value of ρ, yielding the fine-level m-rep M_1. The interpolation and refitting are repeated for a desired number of coarse-to-fine levels. The procedure is illustrated in Fig. 3.3.

The homology between the positions of corresponding medial atoms in m-reps that represent different instances of an object is ensured by constraining the distances between pairs of neighboring medial atoms in the coarse m-rep M_0 to be equal (this constraint is enforced during deformable segmentation). Homology between corresponding atoms in m-reps M_i and M_{i+1} that describe the same instance of the object at different levels of detail is ensured by requiring that during the deformation of the prediction m-rep M_i', its atoms are constrained to move only in the direction orthogonal to the medial axis, with the exception of end atoms, which are allowed to move freely. This constraint ensures that the statistical features computed by comparing corresponding atoms in M_i' and M_{i+1} measure geometrical differences between the two representations of an object at different scales and do not reflect the movement of non-terminal medial atoms in the direction along the medial axis, which is irrelevant for describing shape differences.

The following section describes in detail the interpolation procedure used to compute a continuous medial locus from a set of medial atoms. Readers interested in the statistical aspects of the coarse-to-fine shape characterization approach may skip over to the subsequent section.

3.2.1. Medial Axis Interpolation. The two-dimensional medial interpolation problem can be stated as

> Given two medial atoms $\mathbf{m}_1$ and $\mathbf{m}_2$, connect them using a continuous locus of medial atoms.

A plausible approach to this problem would be to express the continuous medial locus by interpolating the positions and radii of $\mathbf{m}_1$ and $\mathbf{m}_2$. Such an interpolation would have to obey a set of constraints over the entire continuous locus and satisfy

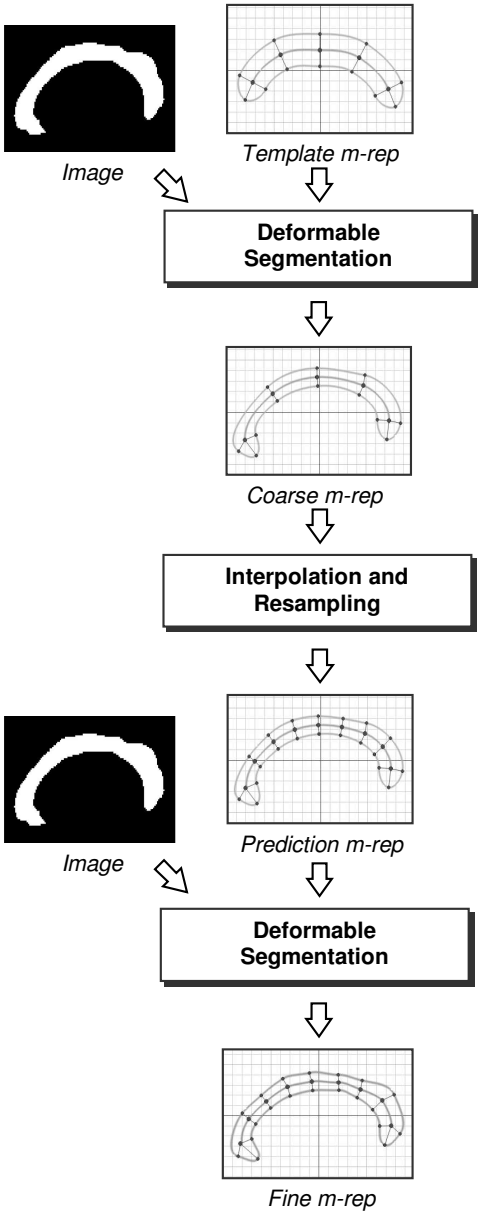


FIGURE 3.3. Outline of the procedure used to fit a coarse-to-fine hierarchy of m-reps to an image. First a *template m-rep* with few atoms is deformed to fit a binary image with large tolerance. The resulting *coarse m-rep* is interpolated and resampled, forming a *prediction m-rep*, which is in turn deformed to fit the same image with lower tolerance, thus forming a *fine m-rep*.

first order boundary conditions imposed by the object angles and axial angles of $\mathbf{m}_1$ and $\mathbf{m}_2$. The constraints and boundary conditions can be expressed in the following form. Let $x(s), y(s), r(s)$ be a medial locus connecting $\mathbf{m}_1$ and $\mathbf{m}_2$, and let this locus be parameterized by the arclength s . Define $\alpha(s) = \arccos \frac{dx}{ds}$, and $\theta(s) = -\arcsin \frac{dr}{ds}$; these two quantities correspond to the frame angle and object angle in a medial atom. The locus must satisfy the boundary conditions imposed by $\mathbf{m}_1$ and $\mathbf{m}_2$ on $x(s), y(s), r(s), \alpha(s)$, and $\theta(s)$, as well as the following constraints:

$$r > 0, \quad \frac{dr}{ds} > 0, \quad \frac{\sin\theta}{r} \neq \frac{d\alpha}{ds} \pm \frac{d\theta}{ds} \ , \tag{21}$$

which follow from [Blum and Nagel 1978] and state that (1) the radius is always positive, (2) the object angle is defined everywhere, and (3) the radius of curvature of the left and right implied boundaries is non-zero. A satisfactory solution to this interpolation problem has not yet been found.

An alternative approach to medial atom interpolation is taken in this chapter. This approach interpolates the boundary nodes of $\mathbf{m}_1$ and $\mathbf{m}_2$ and then finds the symmetry set of the interpolated boundary segments. A first order interpolation of the boundary nodes is constructed using cubic Hermite splines (this is identical to the interpolation employed to construct a continuous m-rep boundary used during segmentation). A pair of medial atoms and Hermite boundary curves fitted to their boundary nodes are shown in Fig. 3.4.

The following procedure is used to compute the symmetry set of the Hermite boundary curves. Let $\mathbf{x}_L$ and $\mathbf{x}_R$ denote the left and right boundary curves, parameterized by arclengths s and t respectively. A bitangent circle, and hence a medial atom, can be fitted at a pair of boundary locations $\mathbf{x}_L(s)$ and $\mathbf{x}_R(t)$ iff

$$F(s,t) = [\mathbf{x}_R(t) - \mathbf{x}_L(s)]^T \left[\frac{\mathrm{d}\mathbf{x}_R(t)}{\mathrm{d}t} - \frac{\mathrm{d}\mathbf{x}_L(s)}{\mathrm{d}s} \right] = 0. \tag{22}$$

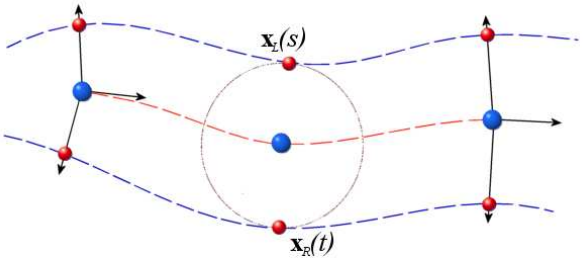


FIGURE 3.4. Medial axis interpolation. The symmetry set is found by solving an implicit equation that holds when a circle can be fitted tangent to two points on the boundary.

The symmetry set is computed by finding the locus of pairs $\{s, t\}$ that satisfy (22) and constructing a medial atom for each pair. In a brute force approach, the locus is found by using Brent's numerical root finder [Presse et al. 1988] to compute s for given values of t. This symmetry set necessarily passes through $\mathbf{m}_1$ and $\mathbf{m}_2$ because the boundary curves are tangent to the circles defined by the positions and radii of $\mathbf{m}_1$ and $\mathbf{m}_2$. The symmetry set is a superset of the Blum medial locus that may include bitangent circles that cross one of the Hermite boundary curves. The Blum medial locus can be found by checking validity of each point of the symmetry set and discarding the invalid points.

This medial interpolation approach does not yield valid medial loci (i.e., loci for which the constraints in (21) are satisfied) for all configurations of medial atoms $\mathbf{m}_1$ and $\mathbf{m}_2$. Some configurations produce interpolated boundaries that cross or fold. If at any point on the locus one of the elements in the vector ∇F is negative, the symmetry set experiences *retrograde motion*: a step forward along the symmetry set results in a step backwards along one or both of the boundaries.

Retrograde motion can occur inside the object formed by the two boundaries if one of the boundary curves contains a codon[1] with a positive curvature maximum and thus induces branching of the medial axis according to [Leyton 1987]. At branches the symmetry set of an object forms a swallowtail, and the medial axis has C_0 continuity. The

[1]A codon is a portion of a curve contained between two minima of curvature

medial axis can still be traced successfully, by removing the points where ∇F has a negative component.

Special difficulty occurs when ∇F has a negative component at one of the two end atoms. In such a case the medial atom does not fall on the Blum medial locus of the Hermite boundary curves, and tracing of the medial locus is not possible. At a medial atom ∇F is determined by the position, tangent and curvature at the implied boundary points. If the medial atom were to include a second order component, as previously suggested, it would be possible to find illegal configurations of medial atoms regardless of the function used to interpolate the boundary.

The corpus callosum data set examined in this chapter does not pose serious problems in terms of medial interpolation. However, when the method is extended to 3D, problems occur more often, since the zero set of F is now a two dimensional manifold in four dimensional space. Even for simple m-reps one encounters problems of retrograde motion. The difficulties of the medial atom interpolation problem has led to the eventual development of continuous m-reps, which are described in Chap. 5.

3.3. Statistical Analysis of Shape: Methods and Results

This section presents a statistical shape analysis approach whose goal is not only to describe shape variability but to provide separate descriptions in terms of geometrical properties (bending or growth), scale (coarse or fine), and location (the whole object or an area). This goal is achieved by constructing a set of medially derived features that are separable in these terms and are invariant under the similarity transform. Application of principal component analysis to these features allows different types of variability to be visualized. Classification techniques make it possible to measure how significantly these different aspects of the description contribute to discrimination between classes of objects. Figure 3.5 summarizes the statistical shape analysis approach.

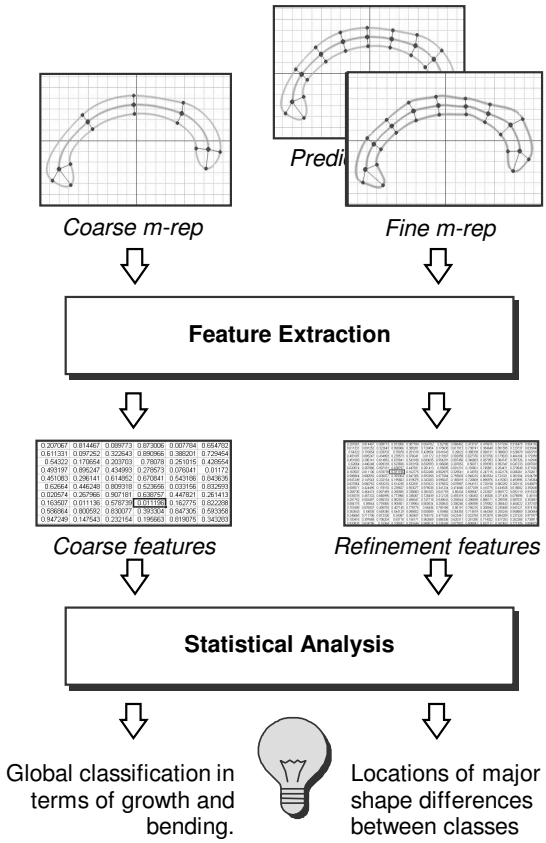


FIGURE 3.5. Computation of features and coarse-to-fine statistical analysis. M-reps at each level of detail are converted to shape features which have common units and are invariant under the similarity transform. Coarse features relate neighboring atoms in a coarse m-rep and refinement features describe local differences between corresponding atoms in the fine and prediction m-reps.

This section presents results of experiments conducted on a data set that consists of binary segmented images of the corpus callosum of 71 subjects in a schizophrenia study (one such image is shown in Fig. 3.2a). The images are organized into two classes corresponding to 40 schizophrenia patients and 31 healthy control subjects. The corpus callosum was selected because its shape can be accurately approximated with a single-figure m-rep and because it has been a subject of shape characterization study by other authors in the field [Golland et al. 1999].

The coarse-to-fine m-rep fitting procedure outlined in the end of Sec. 3.2 was applied to the 71 corpus callosum images. The set of levels of detail at which to represent these objects was chosen empirically by examining different options. Due to the rather simple shape of the corpus callosum, just two levels of detail were used; i.e., each corpus callosum was described by a coarse m-rep M_0, a prediction m-rep M_0', and a fine m-rep M_1. The coarse m-reps were made to consist of five medial atoms, as any smaller number of atoms did not provide enough flexibility to capture the shape of the corpus callosum, even at a very coarse scale. The fine m-reps consist of nine medial atoms, that is, one atom was inserted between pairs of adjacent atoms. The values of the aperture-to-width ratio ρ were also chosen empirically. By way of visual observation the value of $\rho = 0.25$ was chosen for fine m-reps. In most cases, this value yielded m-reps that describe the boundary of the corpus callosum with a sub-pixel accuracy, with the exception of the its, where the Blum medial locus of the corpus callosum boundary often contains small branches, and the single-figure m-rep sometimes can not be fitted accurately regardless of the level of detail used. For coarse m-reps, the value of $\rho = 1$ was used. It was chosen after observing that the choice $\rho = 0.5$, which is proportional to the sampling rate of the coarse m-reps, did not result in significantly large differences between coarse and fine m-reps. Examples of the m-reps used to represent the corpus callosum are shown in Fig. 3.2.

3.3.1. Statistical Features Derived from M-Reps. The geometric form of an object is defined as the set of geometric properties that are not affected by translating, rotating, and scaling the object [Kendall 1984]. Subsequently, statistical features describing shape must be invariant under the similarity transform. Therefore, the parameters of medial atoms, which include positions, widths, and orientations, can not be directly used as shape features.

In order to produce features that do not reflect the global position, orientation, and size of each object, shape analysis methods commonly use an alignment procedure, such as the Procrustes algorithm [Gower 1975, Cootes et al. 1994, Dryden and Mardia 1998]. When such procedures are used, local measurements of variability can become "polluted" by global shape variability because each point's location relative to the corresponding points on other objects is influenced by the locations of all the other points in the object. For example, consider aligning two m-reps that are nearly identical, except for the i-th medial atom, whose position is very different between the two m-reps. Procrustes alignment, which minimizes the mean square distance between corresponding points, would align the objects in such a way that pairs of atoms that were originally coincident would have different positions, while the disparity in the position of the i-th atom would be reduced. Thus the alignment algorithm would propagate the large difference in the i-th atom to all the other atoms, making it more difficult to detect that the difference between the two m-reps is only exhibited locally by the i-th atom.

As an alternative to alignment, the approach taken in this chapter bases statistical features on unit free measurements that describe relationships between medial atoms in m-reps. These features are computed differently for the most coarse m-rep M_0 and for the finer m-reps $M_1, M_2, \ldots$.

Features that describe the relationships between medial atoms in M_0 are called *coarse features*. Five types of coarse features are computed; some of them capture geometrical properties that describe the local bending of objects, while others describe local object

width and elongation. The definitions of the features are given in Table 3.1. Feature F_i^1 describes how the object's medial locus bends at a medial atom and F_i^2 describes the bending of the medial axis between atoms. Feature F_i^3 captures the width of the object at a medial atom, and F_i^4 describes the local change in width at the atom. Feature F_i^5 is only computed at end atoms and measures the elongation of the object's ends. All coarse features are unit free because they describe ratios between different object-related distances, either explicitly or as sines or cosines of angles.

A fine level m-rep M_j includes the undesirable absolute position information, as well as the more global shape information already described by M_{j-1}. When analyzing fine level m-reps, we are interested in answering the question, "What new information does the fine m-rep bring to the shape description?" This question is answered by analyzing the residual differences between the coarser and finer m-reps. Since the m-rep M_{j-1} is equivalent to the same level prediction m-rep M'_{j-1} and since M'_{j-1} and M_j have the same number of components, the difference between M'_{j-1} and M_j is used as the basis for constructing a set of residual features that describe the refinement achieved at level j. These features are called *refinement features*.

TABLE 3.1. Definitions of the features used for the statistical analysis.

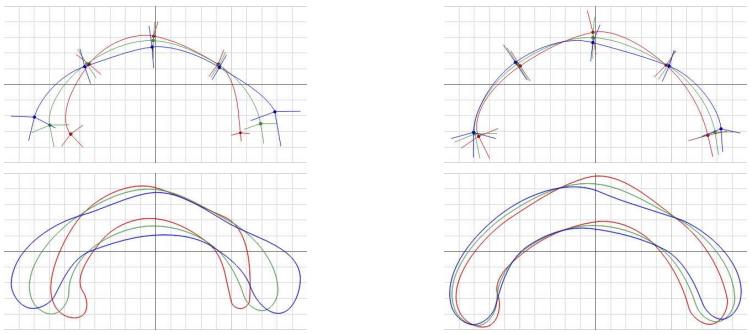
Coarse Feature	Range of i	Refinement Feature	Range of i
$F_i^1 = \sin\angle\vec{l}_i\vec{l}_{i+1}$	$2\ldots n-1$	$G_i^1 = \frac{(\mathbf{x}_i-\mathbf{x}'_i)\cdot\perp\vec{b}'_i}{\|\vec{l}'_i\|}$	$1\ldots n$
$F_i^2 = \sin\angle\vec{b}_i,\vec{l}_i$	$1\ldots n$	$G_i^2 = \frac{(\mathbf{x}_i-\mathbf{x}'_i)\cdot\vec{b}'_i}{\|\vec{l}'_i\|}$	$1, n$
$F_i^3 = \frac{r_i}{\|\vec{l}_i\|}$	$1\ldots n$	$G_i^3 = \sin\alpha'_i - \alpha_i$	$1\ldots n$
$F_i^4 = \cos\theta_i$	$1\ldots n$	$G_i^4 = \sin\theta'_i - \theta_i$	$1\ldots n$
$F_i^5 = \frac{d_i}{\|\vec{l}_i\|}$	$1, n$	$G_i^5 = \frac{r_i-r'_i}{\|\vec{l}'_i\|}$	$1\ldots n$
		$G_i^6 = \frac{d_i-d'_i}{\|\vec{l}'_i\|}$	$1, n$

$$\vec{l}_i = \begin{cases} \mathbf{x}_2 - \mathbf{x}_1 & \text{if } i = 1 \\ \mathbf{x}_i - \mathbf{x}_{i-1} & \text{if } i > 1 \end{cases}$$

Table 3.1 lists six types of unit free refinement features computed using a prediction m-rep M'_{j-1} and a fine m-rep M_j (the subscript j is omitted for brevity and $\mathbf{x}'_i$ is used to refer to the position of the i-th atom in M'_{j-1}, while $\mathbf{x}_i$ refers to the position of the corresponding atom in M_j). The first two refinement features represent the spatial displacement of the fine medial atom with respect the corresponding atom in M'_{j-1}. G_i^1 measures the component of this displacement in the direction normal to the medial axis of M'_{j-1}. For internal medial atoms this component captures all of the displacement since these atoms are not allowed to move along the axis during segmentation. End atoms are allowed to move freely, and G_i^2 is also used to capture their movement. The remaining features measure differences in the orientation of the medial axis (G_i^3), int the object angle (G_i^4), in width (G_i^5), and in end elongation (G_i^6) between corresponding pairs of medial atoms.

These sets of coarse and refinement features were chosen in a way that makes it possible to reconstruct m-reps from their feature representation. Given the absolute location $\mathbf{x}$, orientation θ, and radius r of just one medial atom, the information contained in the coarse features can be used to rebuild the coarse m-rep M_0. Similarly, given a reconstructed m-rep M_{j-1} and the refinement features corresponding to level j, the m-rep M_j can be reconstructed. Since the inter-atom sampling distance is not included as one of the features, reconstruction relies on regular sampling of atoms being enforced during segmentation.

3.3.2. Visualization of Shape Variability. When statistical features are defined in such a way that the geometrical form of an object can be reconstructed from its feature space representation, the variability in a population of objects can be visualized using the procedure outlined in Cootes et al. [1994]. An informative direction in feature space, such as one of the principal modes of variability computed by PCA, can be displayed as an animation whose frames are m-reps reconstructed from a sequence of points in feature

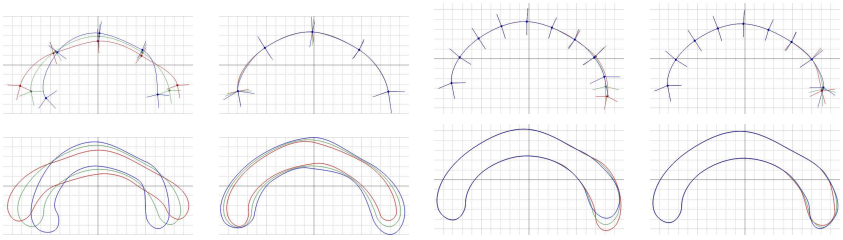


First Mode Second Mode

FIGURE 3.6. Shapes reconstructed along the first two primary modes of variability in coarse features. All coarse features are combined in this figure, without separation into growth and bending. Top row shows the medial axis and the bottom row shows the implied boundary. In each image the colors represent positions along the mode: -2σ in red, mean in green and $+2\sigma$ in blue.

space sampled along this direction. Alignment of the reconstructed m-reps is necessary to produce a visually meaningful animation because the coarse statistical features lack information about the absolute placement of the m-rep. For visualization purposes only, the Generalized Procrustes alignment algorithm [Gower 1975], which was purposely not used to compute features in Sec. 3.3.1, is applied to the locations of the medial atoms belonging to all of the reconstructed coarse m-reps.

The feature space formed by the coarse-to-fine approach consists of subspaces that describe different aspects of shape variability. The subspace formed by the coarse features describes the global coarse-scale shape variability, and subspaces formed by refinement features taken at different regions of the object describe local fine-scale variability. These subspaces can be further divided into parts that describe variability separately in terms of bending, widening, and elongation. PCA can be applied separately to any one of these subspaces, resulting in different informative directions in feature space, which can be visualized. Such visualizations of different aspects of shape variability are shown in Figs. 3.6 and 3.7.



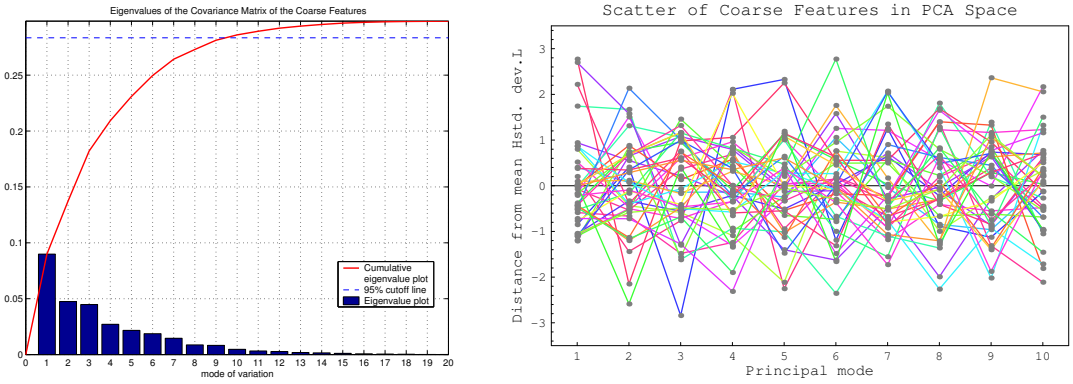
a. b. c. d.

FIGURE 3.7. Modes of variability of corpus callosum shape. a. First mode in coarse-level bending. b. First mode in coarse-level growth, i.e. widening and elongation. c. First local mode of variability at posterior of the corpus callosum. d. Second local mode at the posterior. Shapes at -2, 0, and $+2$ standard deviations along the mode are shown in red, green and blue. Top row shows medial geometry and bottom row shows the implied boundaries.

Figure 3.6 visualizes the principal components of variability in the coarse features computed for the corpora callosa of the 40 schizophrenia patients. Reconstructed m-reps and their implied boundaries are shown for the mean shape and for feature space points sampled ± 2 standard deviations away from the mean along the first two principal modes of variability. We see that the first principal mode of variability captures both the bending of the corpus callosum and the widening of its ends. The second mode largely captures the bending of the right half of the corpus callosum and the widening in the midbody.

Figure 3.8a summarizes the PCA performed on coarse features by plotting the eigenvalues of the covariance matrix and their cumulative sums. This plot shows that $k_{0.95} = 10$ dimensions are needed to capture 95% of the variability. Figure 3.8b shows projections of the coefficient vectors $\vec{\omega}_i$ onto the coordinate directions in coefficient space. Aside from a few exceptions, training shapes lie within two standard deviations from the mean.

The ability to separate variability into localized components and the ability to decompose global variability into bending and growth are strengths of the m-rep based shape characterization method. Figure 3.7a visualizes the primary mode of variability for only the coarse bending features (F_i^1 and F_i^2). It appears that the change in bending



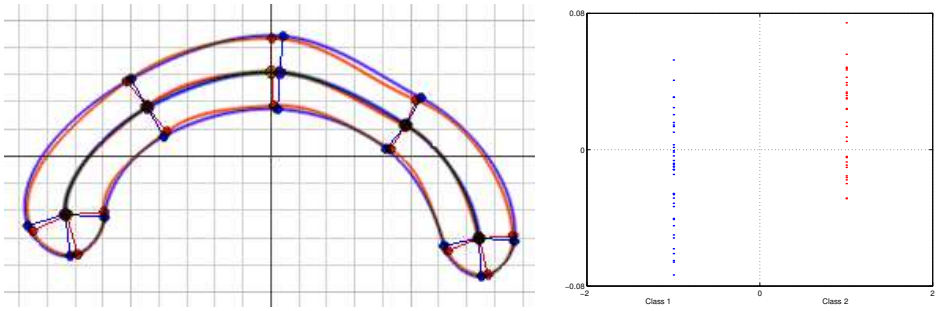
a. b.

FIGURE 3.8. Example of principal component analysis (PCA). a. Eigenvalues of the covariance matrix of features computed from coarse m-reps (schizophrenic corpus callosum). b. Parallel coordinate plot of the distribution of the coarse m-rep distribution in the PCA coefficient space.

is primarily explained by the fact that some corpora callosa are curved like a 'C' while others are more straight. A similar type of bending can be observed in Fig. 3.6a, which shows the first mode of variability of the combined coarse features. Figure 3.7b shows the first mode of variability in coarse widening and elongation features (F_i^3, F_i^4 and F_i^5). The main action observed here is the uniform widening of the corpus callosum.

Figures 3.7(c,d) show local shape variability that occurs at the posterior end of the corpus callosum, i.e., in features $G_i^1, \ldots, G_i^6$. Here the primary mode of variability mostly corresponds to elongation of the tip of the corpus callosum. When PCA is applied to a subset of all the coarse or refinement features, the reconstruction algorithm used for visualization substitutes the sample mean value for the features not included in the PCA. Hence the anterior and the midbody of the reconstructed corpora callosa in Fig. 3.7(c,d) are those of the mean coarse shape.

3.3.3. Simulated Data for Classification Study. The corpus callosum data set is not well suited for classification experiments because no statistically significant shape differences between the class of schizophrenia patients and the class of healthy controls could be found using this chapter's methodology. For example, Fig. 3.9a shows that



a. b.

FIGURE 3.9. A comparison of the schizophrenic corpus callosum shape to the healthy shape on the basis of coarse features. a. An m-rep corresponding to the mean schizophrenic shape (blue) and the m-rep corresponding to the mean healthy shape (red). The mean shapes are very similar. b. Coarse features of schizophrenic and normal corpora callosa projected onto the Fisher linear discriminant.

the means of coarse features describing these two classes are very similar, and Fig. 3.9b shows a significant overlap in the projections of the coarse m-rep features of each member of the schizophrenic class (in blue) and each member of the healthy class (in red) onto the Fisher linear discriminant computed for these two classes. There is little separation in the discriminant direction. Leave-one-out validation of a classifier based on the Fisher linear discriminant constructed on the coarse features yielded the correct classification rate of 0.51, which is equivalent to random guessing. Other researchers have reported similar negative results for corpus callosum discrimination in schizophrenia, e.g., [Highley et al. 1999].

To demonstrate the diagnostic ability of the coarse-to-fine approach in a case that supports discrimination, artificial classes were constructed. Three classes of objects were simulated on the basis of the corpus callosum shape. Representatives of these classes are illustrated in Fig. 3.1. Classes 1 and 2 differ in coarse shape while classes 1 and 3 have the same coarse shape but differ locally: class 3 has a random bump at the midbody of the corpus callosum. The method ought to be able to discriminate between classes 1 and 2 globally while discriminating locally between classes 1 and 3.

The simulated corpus callosum images were synthesized using the elliptical harmonics representation [Székely et al. 1996] of the segmented corpora callosa, which was kindly provided by the group headed by Dr. Gerig. In this representation, an object is described by a set of coefficients yielded by the decomposition of the object's boundary into elliptical harmonics basis functions. This representation was used to synthesize the shape of the coarse corpus callosum coarsely, and in order to do so only the coefficients of the five lowest-frequency elliptical harmonics were used. Two multivariate normal distributions with identity covariance matrices were set up in the space whose basis is formed by the principal modes of the elliptical harmonic coefficients. The means of the distributions were chosen in such a way that the expected probability of correctly classifying a sample from either distribution using a maximum likelihood classifier was equal to 0.80.

Classes 1 and 2 were synthesized by randomly sampling from these two distributions. The differences between these two classes are global (i.e., exhibited over the entire object) and coarse in scale. Class 3 was sampled from the same distribution as Class 1, but an additional randomized deformation in the form of a bump was applied to the boundaries of the objects in Class 3 in order to simulate local shape variability. The bump was generated by pushing the boundary in a selected region in the midbody of the of the corpus callosum outwards along the boundary normal. The shape of the bump was defined using a boundary displacement in the form of a cosine function truncated at $\pm\pi/2$. The position and amplitude of the bump were sampled randomly from a normal distribution.

Boundaries of the three classes were used to create binary images, and coarse and fine m-reps were fitted to these images using the procedure described in Sec. 3.2. M-reps reconstructed from feature space points corresponding to the sample means of the three classes are shown in Fig. 3.10.

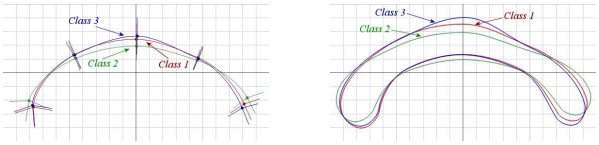


FIGURE 3.10. Mean m-reps of the three simulated classes and their implied boundaries.

TABLE 3.2. Classification using five different feature sets. Shown are type I and II errors for leave-one-out classification of simulated corpus callosum classes 1 vs. 2 and 1 vs. 3. We use the notation $p(C_i|C_j)$ to mean the probability of classifying a member of class C_j incorrectly into class C_i.

	$p(C_2\|C_1)$	$p(C_1\|C_2)$	$p(C_3\|C_1)$	$p(C_1\|C_3)$
All coarse features	0.32	0.28	0.12	0.08
Coarse growth features	0.36	0.24	0.08	0.0
Coarse bending features	0.20	0.36	0.08	0.04
Refinement features at anterior	0.36	0.32	0.36	0.24
Refinement features at midbody	0.36	0.36	0.12	0.04

3.3.4. Classification Results. Classification based on the Fisher linear discriminant was applied to pairs of classes (1, 2) and (1, 3) in leave-one-out cross-validation experiments[2]. Table 3.2 shows the results of these experiments with classification applied to different sets of features: (i) all coarse features, (ii) just the bending coarse features, (iii) just the widening and elongation coarse features, (iv) refinement features at anterior end of the corpus callosum, and (v) refinement features at midbody of the corpus callosum, where the bump in class 3 is located.

Table 3.2 shows that the differences between classes 1 and 2 are detected slightly better by the coarse features than by the local refinement features. This is to be expected because shapes in these two classes come from different distributions in elliptical harmonics coefficient space and have relatively less fine-scale difference in boundary detail. The bending features discriminate between classes 1 and 2 better than other coarse

[2]Since the objects in classes 1, 2, and 3 were generated synthetically, one may ask why the decision to use leave-one-out cross-validation for testing was made, and not a second separate set of objects generated. The reason for this decision was the time cost of fitting m-reps to these objects using a procedure that requires some user intervention.

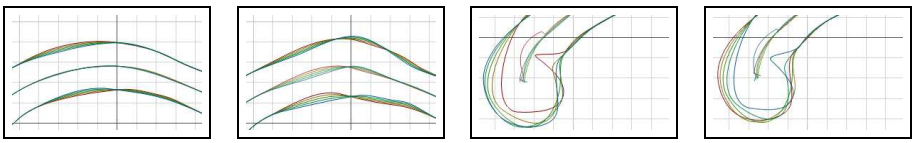


FIGURE 3.11. Local variability in primary mode of refinement features in classes 1 and 3 at midbody and anterior of the corpus callosum. Shapes at $-2, -1, 0, +1, +2$ standard deviations are shown. This plot is a 'zoomed in' analog of the plots in Fig 3.7 c,d, which show local shape variability in the corpus callosum of schizophrenia patients.

features, suggesting that global shape differences mostly exhibited in the bending of the corpus callosum. Classification between classes 1 and 3 based only on the coarse features yields high accuracy because the coarse description is sensitive enough to detect and be affected by the bump. However, the localized classification based on refinement features helps detect the presence of the bump and find its location.

In order to compare the significance of inter-class shape difference in different locations in the corpus callosum, the object was divided into overlapping neighborhoods. A neighborhood was defined as a pair of adjacent medial atoms in the fine m-rep model. The sets of refinement features generated by each of these neighborhoods were analyzed separately. Fig. 3.12 shows the measure of discriminability between pairs of classes in each of these neighborhoods. Discriminability of a neighborhood is computed in this case as the p-value of a multivariate Fisher T-test applied to the features that describe the neighborhood. A relative comparison of the values leads to a conclusion that a relatively significant shape difference between classes 1 and 3 is found near the middle of the object, while classes 1 and 2 do not exhibit such a difference. Hence, by measuring the relative difference in discriminability between classes in each neighborhood and finding the neighborhood with the greatest difference, the bump was detected and localized successfully.

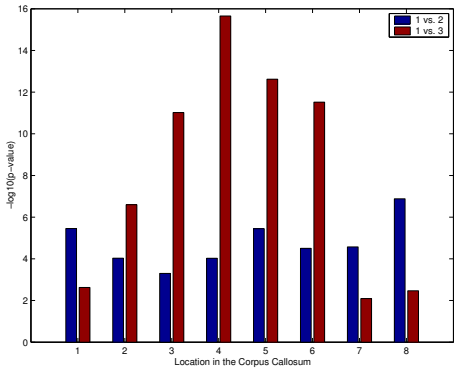


FIGURE 3.12. Discriminability between classes 1 vs. 2 and 1 vs. 3 at each location in the corpus callosum. The ordinate values are negated exponents of the p-value of Fisher T-test.

3.4. Discussion

The major contribution of this chapter is the development and partial evaluation of a hierarchical method of shape analysis that reduces the dimensionality of shape characterization by dividing the problem into a set of subproblems that examine objects at different scales. Future work will show whether such coarse-to-fine decomposition of the problem improves shape analysis of clinical data.

The coarse-to-fine approach decomposes the shape characterization problem into subproblems of lower dimensionality. In the process it makes an assumption that the shape variability described separately by the parts (subproblems) sufficiently describes the variability exhibited by the whole of the parts. The hierarchical decomposition into subproblems incurs a risk that correlational information between parts of the fine-scale representation of the object may be lost. For instance, if an object has two fine-scale bumps that are located in different regions, and whose sizes are correlated, even a hierarchical method based on Markov random fields [Lu et al. 2003a] might not be able to detect this correlation, while a method that works with the entire fine-scale representation of the object may. The effect of this decoupling of correlation on real shape problems remains

to be studied, as does the question of explicit modeling of the correlations between the subproblems.

An extensive application of the coarse-to-fine methodology to clinical shape studies would require the methodology to be extended to handle three-dimensional objects and objects that require m-reps with more than a single figure for an accurate medial description. Recent progress in m-rep segmentation has made it possible to automatically fit m-rep templates to collections of images of 3D anatomical structures [Pizer et al. 2000, Joshi et al. 2001]. Even more recent research by Fletcher et al. [2003b,a] on principal geodesic analysis promises to produce effective and mathematically founded feature sets for use in m-rep based 3D shape analysis.

The major theoretical difficulties involved in extending the approach to 3D are tied to the problem of medial interpolation and to the problem of maintaining correspondences between m-rep descriptions of different objects. The interpolation problem is addressed by Chap. 5, where a new flavor of continuous m-reps is developed. These m-reps by construction allow arbitrary sampling and are well fitted for the needs of the a more flexible coarse-to-fine approach.

To be ultimately useful in practice, our method can not be limited to single figure objects. Few shapes can be accurately represented by a single chain of medial atoms or, in 3D, by a single mesh. The capability to analyze multi-figural objects can be achieved easily if the medial branching topology is the same for all shapes in the training set. This capability requires small scale residues, e.g., via the methods of Pizer et al. [2003] that add the boundary level of detail in terms of displacements along the normal vectors. In this case, one can simply formulate new features that describe figure-to-figure relationships and the boundary displacements. The method can even be extended to data that includes multiple objects, hopefully addressing problems as challenging as the shape analysis of the whole abdomen.

Fixed multi-figural medial branching topology is an improvement over single figure topology in terms of accuracy of the boundary approximation. However, no implementation with fixed topology can accurately and fully capture certain sets of real-world shapes, especially at the fine scale. Since the coarse-to-fine approach establishes and uses rigid correspondences between medial atoms across a set of shapes, major theoretical advances are required to extend the approach to handle variability in m-rep topology.

The addition of the *boundary level of detail* will improve the accuracy of fixed topology analysis. This new level captures the difference between a boundary implied by a fine-scale m-rep and the actual object boundary in terms of displacements along the normal vectors. Shapes with different medial topology can be approximated by a combination of the boundary level of detail and a medial representation with fixed representation.

The problem of correspondence, described in Sec. 2.2.4 is common to nearly all boundary and medial based shape analysis approaches, including the one described in this chapter, where homology is established by forcing medial atoms to be equally spaced in each object. Such correspondence is highly sensitive to the placement of the end atoms of m-reps. Better correspondence could be achieved by incorporating the statistical properties of the training set, along the lines of the Davies et al. [2002] minimal description length method.

CHAPTER 4

Feature Selection for Shape-Based Classification of Biological Objects

In this chapter, feature selection methodology from the machine learning literature is adapted and applied to the problem of statistical shape-based classification of biological objects. The feature selection paradigm is used to discover the regions of objects where the difference between classes is most pronounced. Feature selection can improve the generalization ability and statistical significance of shape-based analysis. A feature selection algorithm based on support vector machines is extended to take advantage of special relationships between neighboring features, which are inherently present in geometric object representations. Performance analysis using simulated and clinical data is presented.

The research described in this chapter was done in collaboration with Dr. Sarang Joshi and Dr. Stephen M. Pizer at the University of North Carolina and with Dr. John G. Csernansky and Dr. Lei E. Wang at the Washington University Medical School, who together contributed to every facet of the work. The results reported below were made possible thanks to the contribution of Dr. Csernansky and Dr. Wang, who provided the clinical data set used to study the effect of schizophrenia on the hippocampus, and by Dr. Guido Gerig, who provided the corpus callosum data for use in synthetic data experiments.

4.1. Introduction

This chapter explores the ways in which a machine learning technique called *feature selection* can be used to improve shape characterization. Feature selection reduces the dimensionality of classification problems by finding the subset of features that best captures the differences between classes. Classifiers restricted to the selected subset of features are less affected by sampling noise and tend to generalize better than the classifiers trained on the entire feature set. Feature selection has been shown to dramatically improve the generalization ability of classifiers in high-dimensional low sample size problems [Bradley and Mangasarian 1998, Weston et al. 2001].

The potential benefit of using feature selection algorithms in shape characterization problems extends beyond the improved generalization ability. Examination of features deemed most relevant by such algorithms may reveal the areas of organs that are most affected by a disease, leading to improved localization and understanding of the biological processes responsible for the disease.

Feature selection algorithms in the machine learning literature usually address general classification problems and make minimal assumptions as to where the different features come from and how they may be related. In shape classification problems, where features are usually derived from dense geometrical object representations, there exist special relationships between neighboring features. By incorporating our knowledge of these relationships into feature selection algorithms, we can improve their performance and stability when applied to shape classification. The two properties of shape features that are particularly useful for improving feature selection are *structure* and *locality*.

We use the term *structure* to indicate the importance of the order in which the features are arranged in a classification problem. In many problems, the order of the features is arbitrary, as is the case, for example, when all the features describe different physical properties of an object, such as its height, weight, age or density. However,

when the features are measurements regularly sampled from a lattice, as is the case in many geometrical object representations, the order of the features is important, as nearby features are more likely to be correlated than the far-away features.

A biological process responsible for variability in the shape of an anatomical object exhibits *locality* if it affects the object at one or at most a few locations, which are consistent across the population of objects. In reference to a feature set, I use the term locality to mean that some components of the statistical variability in the data can be localized to one or more subsets of the features.

In the absence of structure and locality, the feature selection problem is purely combinatorial, since in the set of n features there are 2^n possible subsets and all of them are considered *a priori* to be equally worthy candidates for feature selection. The properties of structure and locality constitute prior knowledge about the kinds of feature subsets that ought to be selected. Feature sets consisting of one or a few contiguous subsets are more likely candidates than feature sets in which the selected features appear scattered. By assuming that shape features exhibit structure and locality, we can reduce the number of possible solutions of a feature selection algorithm.

The feature selection method presented in this chapter is best suited for the shape analysis approaches that use dense object representations. This rich class of representations includes boundary point distribution models [Cootes et al. 1994], parametric boundary models based on Fourier and spherical harmonic basis decomposition [Staib and Duncan 1992, Kelemen et al. 1999], discrete and continuous medial representations [Pizer et al. 1999, Yushkevich et al. 2002, Golland et al. 1999], as well as functional object representations, such as distance transforms [Golland et al. 2001], and deformation fields based on a warping of a template to each object in the training set [Csernansky et al. 1998, Davatzikos et al. 1995, Joshi et al. 1997]. The features yielded by these representations exhibit structure and locality because they are densely and regularly sampled.

Many of the methods in statistical shape analysis focus on estimating the probability distribution on the shape space [Cootes et al. 1994, Kelemen et al. 1999, Staib and Duncan 1992, Joshi et al. 1997]. Such probability distributions can be used as priors for deformable segmentation and can be sampled to visualize shape variability. Principal component analysis, which is used by many of the methods to estimate the shape distribution, is related to feature selection, as it reduces the dimensionality of the data to a linear combination of the original features.

While the issues of representation and correspondence have been the focus of extensive research in the shape characterization literature, the application of the feature selection paradigm to the shape characterization problem has received less attention. However, there has been considerable work in the literature in using classification techniques to detect, localize and describe the anatomical differences in the shape between different populations [Golland et al. 2002, Gerig et al. 2001, Styner 2001, Csernansky et al. 1998, Joshi et al. 1997, 2002, Yushkevich et al. 2001].

The chapter is organized as follows. Section 4.2 presents the methodology that incorporates prior knowledge about structure and locality of shape features into an existing feature selection method by Bradley and Mangasarian [1998]. Section 4.3.1 studies the performance of the algorithm in simulated data examples, where the features are normally distributed and the relevant features are localized. In Section 4.3.2 the method is applied to a simulated shape classification problem. Section 4.4 shows the results of applying the method to real clinical data.

4.2. Methods

This section contains the details of the new feature selection method for shape classification, which is an extension of an existing feature selection algorithm by Bradley and Mangasarian [1998]. The novelty of our method is that it searches for an optimal set of *windows* of features, as described in the following sections. For clarity, I will use the

term *feature selection* to refer to the original algorithm, while referring to our extended version using the term *window selection*.

This section is organized as follows. Subsection 4.2.1 formulates the feature selection problem in general terms of energy minimization. Subsection 4.2.2 expands this formulation by adding an energy term that favors feature sets that exhibit locality and structure. The concept of minimal window cover is used to measure how localized and structured a feature set is. Subsection 4.2.3 describes how the feature window selection problem can be formulated and solved using linear programming.

4.2.1. Feature Selection. Feature selection is a machine learning methodology that reduces the number of statistical features in high-dimensional classification problems by finding subsets of features that are most relevant for discrimination. Classifiers constructed in the subspace of the selected features tend to generalize to new data better than classifiers trained on the entire feature set.

Feature selection methods fall into categories of *filter methods*, which use feature selection as a preprocessing step to classification (e.g. [Kira and Rendell 1992]), and *wrapper methods*, which use classification internally as a means of selecting features (e.g. [John et al. 1994, Kohavi and John 1997, Jebara and Jaakkola 2000, Weston et al. 2001]).

This paper uses and extends a wrapper method developed by Bradley and Mangasarian [1998] and Bradley et al. [1998]. This method uses elements from support vector machine theory and formulates feature selection as a smooth optimization problem, which can be transformed into a sequence of linear programming problems.

The input to the feature selection algorithm consists of a training set of objects that fall into two classes, of sizes m and k. Each object is represented by an n-dimensional feature vector. The classes are represented by the feature matrices $\mathbf{A}_{m\times n}$ and $\mathbf{B}_{k\times n}$.

We wish to find the set of features, i.e., a subset of columns of $\mathbf{A}$ and $\mathbf{B}$, that are most relevant for discriminating between the two classes. The idea of Bradley and Mangasarian

[1998] is to look for a relevant subset of features by finding a hyperplane

$$P = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{w}^T \mathbf{x} = \gamma \right\} \tag{23}$$

that optimally separates the two classes, while lying in the minimal number of dimensions, as formulated by the energy minimization problem

$$P = \arg\min_{\gamma, \mathbf{w}} E_{\text{sep}}(\gamma, \mathbf{w}) + \lambda E_{\text{dim}}(\mathbf{w}) \ . \tag{24}$$

The term E_{sep} measures how well the hyperplane P separates the elements in $\mathbf{A}$ from the ones in $\mathbf{B}$. It is expressed as

$$E_{\text{sep}}(\gamma, \mathbf{w}) = \frac{1}{m} \left\| (-\mathbf{A}\mathbf{w} + \mathbf{e}\gamma + \mathbf{e})_+ \right\|_1 + \frac{1}{k} \left\| (\mathbf{B}\mathbf{w} - \mathbf{e}\gamma + \mathbf{e})_+ \right\|_1 \tag{25}$$

where $\mathbf{e}$ represents a vector of appropriate size whose elements are all equal to 1, and $(\bullet)_+$ is the operation that replaces the negative elements of $\bullet$ with zero.

Let P^- and P^+ be a pair of hyperplanes parallel to P, whose distance to P is $1/\|\mathbf{w}\|$. Then, E_{sep} measures the distance to P^+ of those elements of $\mathbf{A}$ that lie on the 'wrong side' of P^+, as well as the distance to P^- of the elements of $\mathbf{B}$ that lie on the 'wrong side' of P^-. By wrong side, I mean that half-space of P^- or P^+ which contains the hyperplane P.

The energy term E_{dim} in (24) is used to reduce the number of dimensions in which the hyperplane P lies. It has the general form

$$E_{\text{dim}}(\mathbf{w}) = \mathbf{e}^T I(\mathbf{w}), \tag{26}$$

where $I(\mathbf{w})$ is an indicator function that replaces each non-zero element of $\mathbf{w}$ with 1. However, since indicator functions are inherently combinatorial and badly suited for optimization, Bradley and Mangasarian suggest approximating the indicator function

with a smooth function

$$I\left(\{w_1 \dots w_n)\}\right) = \left\{1 - \varepsilon^{-\alpha|w_1|}, \dots, 1 - \varepsilon^{-\alpha|w_n|}\right\}, \tag{27}$$

which, according to [Bradley et al. 1997], yields the same solutions as the binary indicator function for finite values of the constant α .

4.2.2. Window Selection for Shape Features. General feature selection algorithms make minimal assumptions about the nature and the properties of features. For instance, the same algorithm may be used for classifying documents on the basis of word frequency or for breast cancer diagnosis. Without prior knowledge of feature properties, the feature selection problem is purely combinatorial, since in a set of n features there are 2^n possible subsets and all of them are considered to be equally worthy candidates for selection.

In shape classification problems, features are typically derived from dense geometrical object representations [Cootes et al. 1994, Staib and Duncan 1992, Kelemen et al. 1999, Pizer et al. 1999, Golland et al. 1999, 2002, Davatzikos et al. 1995, Joshi et al. 1997], and special relationships exist between features derived from neighboring locations in the objects. We hypothesize that by incorporating the heuristic knowledge of these relationships into a feature selection algorithm, we can improve its performance and stability when applied to shape classification.

Features that describe shape are geometric in nature, and the concept of distance between two features can be defined. Furthermore, natural biological processes exhibit *locality*: geometric features capturing shape of anatomical objects that are close together are likely to be highly correlated. General features, such as word frequencies in documents, may not exhibit this property of locality.

Locality makes it possible to impose a prior probability on the search space of a feature selection algorithm. Locality implies that feature sets consisting of one or a few clusters are more likely candidates than feature sets in which the selected features are

isolated. To reward locality, the energy minimization in (24) is expanded to include an additional term:

$$P = \arg\min_{\gamma,\mathbf{w}} E_{\text{sep}}(\gamma, \mathbf{w}) + \lambda E_{\text{dim}}(\mathbf{w}) + \eta E_{\text{loc}}(\mathbf{w}) \ . \tag{28}$$

The term $E_{\text{loc}}(\mathbf{w})$ rewards selection of neighboring features by favoring vectors $\mathbf{w}$ whose non-zero elements form a small number of clusters.

Let $J \subset \{1 \dots n\}$ be the set of features for which $\mathbf{w}$ is non-zero. To measure how clustered the components of J are, we define an 'alphabet' of structured subsets of $\{1 \dots n\}$ called *windows* and measure the most compact description needed to express J using this alphabet.

The neighborhood relationships between the features in the set $\{1 \dots n\}$ depend on the structure of the space from which the features are sampled. Typically, as in the case of parametric shape descriptions, the underlying structure of a feature set is a lattice of one or two dimensions.

In order to define an alphabet of windows over the feature set, we use a metric $d(i, j)$ that assigns a non-negative distance to every pair of features i, j. A set $W \subset \{1 \dots n\}$ is defined to be a *window of size* q if (i) $d(i, j) \leq q$ for all $i, j \in W$, and (ii), there does not exist a superset of W in $\{1 \dots n\}$ for which the condition (i) holds.

The distance function allows us to define windows on arbitrarily organized features. For instance, when features are organized in a one-dimensional lattice, and the distance function is $d(i, j) = |i - j|$, the windows are contiguous subsets of features. By letting $d(i, j) = |i - j| \mod n$, one can allow for wrap-around windows, which are useful for periodic features, such as features sampled along the boundary of a closed object. On higher-dimensional lattices, different distance functions such as Euclidean distance and Manhattan distance generate differently shaped windows. For features sampled from

vertices on a mesh, windows can be constructed using the transitive distance function, which counts the smallest number of edges on a mesh that separate a pair of vertices.

Let $\mathbf{W} = \{W_1 \dots W_N\}$ be a set of windows of various sizes over the feature set $\{1 \dots \mathrm{n}\}$. A *minimal window cover* of a feature subset J is defined as a minimal set $\alpha \subset \{1 \dots N\}$ for which

$$J = \bigcup_{i \in \alpha} W_i \ . \tag{29}$$

We take the locality energy component $E_{\mathrm{loc}}(\mathbf{w})$ to be equal to the size of the minimal window cover of the set of non-zero features in the vector $\mathbf{w}$. While such a formulation is combinatorial in nature, in the following sections we express it in terms of linear programming and derive an elegant implementation.

4.2.3. Linear Programming Formulation. Bradley and Mangasarian express the energy minimization problem in (24) as a series of linear programming problems [Bradley and Mangasarian 1998]. This section briefly summarizes their approach and extends it to include the formulation (28).

The term $E_{\mathrm{sep}}(\mathbf{w}, \gamma)$ in (25) is linear. The global minimum of E_{sep} can be found by solving the following linear programming problem:

$$\begin{aligned} \underset{\gamma, \mathbf{w}, \mathbf{y}, \mathbf{z}}{\text{minimize}} \quad & \tfrac{\mathbf{e}^T \mathbf{y}}{m} + \tfrac{\mathbf{e}^T \mathbf{z}}{k} \ , \\ \text{subject to} \quad & -\mathbf{A}\mathbf{w} + \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{y} \\ & \mathbf{B}\mathbf{w} - \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{z} \\ & \mathbf{y} \geq 0, \ \mathbf{z} \geq 0 \ . \end{aligned} \tag{30}$$

The feature selection problem in (24) can be formulated as the following smooth non-linear program: [1]

$$
\begin{array}{ll}
\underset{\gamma,\mathbf{w},\mathbf{y},\mathbf{z},\mathbf{v}}{\text{minimize}} & \frac{\mathbf{e}^T\mathbf{y}}{m} + \frac{\mathbf{e}^T\mathbf{z}}{k} + \lambda \mathbf{e}^T I(\mathbf{v}), \\
\text{subject to} & -\mathbf{A}\mathbf{w} + \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{y} \\
& \mathbf{B}\mathbf{w} - \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{z} \\
& \mathbf{y} \geq 0,\ \mathbf{z} \geq 0\ , \\
& -\mathbf{v} \leq \mathbf{w} \leq \mathbf{v}\ .
\end{array}
\tag{31}
$$

This formulation does not directly minimize the objective function (24), but rather it minimizes positive vectors $\mathbf{y}, \mathbf{z}$, and $\mathbf{v}$ that constrain the components of the objective function. Such a transformation of the minimization problem is frequently used in support vector methodology in order to apply linear or quadratic programming to energy minimization problems.

The vector $\mathbf{v}$ constrains $\mathbf{w}$ from above and below and thus eliminates the need for using the absolute value of $\mathbf{w}$ in the objective function, as is done in (25). The non-zero elements of $\mathbf{v}$ correspond to selected features.

In order to introduce the locality energy E_{loc} into the linear program, we can express the non-zero elements of $\mathbf{v}$ as a union of a small number of windows and penalize the number of windows used. Let $W_1 \ldots W_N$ be an 'alphabet' of windows, as defined in Sec. 4.2.2. Let $\mathbf{\Omega}$ be an $n \times N$ matrix whose elements ω_{ij} are equal to 1 if the feature i belongs to the window W_j, and are equal to 0 otherwise. Let $\mathbf{u}$ be a sparse positive vector of length N whose non-zero elements indicate a set of selected windows. Then the non-zero elements of $\mathbf{\Omega u}$ indicate a set of features that belong to the union of the windows selected by $\mathbf{u}$.

In order to implement window selection as a smooth non-linear program, the terms $\mathbf{u}$ and $\mathbf{\Omega u}$ are used in place of $\mathbf{v}$ in the objective function. The resulting formulation

[1] Bradley and Mangasarian [1998] call this problem *Feature Selection Concave (FSV)*.

penalizes both the number of selected windows and the number of features contained in those windows:

$$
\begin{aligned}
\underset{\gamma,\mathbf{w},\mathbf{y},\mathbf{z},\mathbf{u}}{\text{minimize}} \quad & \tfrac{\mathbf{e}^T\mathbf{y}}{m} + \tfrac{\mathbf{e}^T\mathbf{z}}{k} + \left(\lambda \mathbf{e}^T \boldsymbol{\Omega} + \eta \mathbf{e}^T\right) I(\mathbf{u}), \\
& -\mathbf{A}\mathbf{w} + \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{y} \\
\text{subject to} \quad & \mathbf{B}\mathbf{w} - \mathbf{e}\gamma + \mathbf{e} \leq \mathbf{z} \\
& \mathbf{y} \geq 0,\ \mathbf{z} \geq 0\ , \\
& -\boldsymbol{\Omega}\mathbf{u} \leq \mathbf{w} \leq \boldsymbol{\Omega}\mathbf{u}\ .
\end{aligned}
\tag{32}
$$

This formulation of the objective function is identical to the energy minimization formulation (28) if none of the windows selected by $\mathbf{u}$ overlap. In case of an overlap, the penalty assessed on the combined number of features in all of the selected windows, and not on the total number of windows in the vector $\mathbf{w}$.

We use the fast successive linear approximation algorithm outlined in [Bradley and Mangasarian 1998] to solve the program (32). The algorithm is randomly initialized and iteratively solves a linear programming problem in which the concave term $I(\mathbf{u})$ is approximated using the Taylor series expansion. The algorithm does not guarantee a global optimum but does converge to a minimum after several iterations. The resulting vector $\mathbf{u}$, whose non-zero elements indicate the selected windows, is very sparse. The *Sequential Object-Oriented Simplex Class Library (SoPlex)*, developed by Wunderling [1996], is used for solving the linear programming problems.

The parameters λ and η affect the numbers of features and windows selected by the window selection algorithm. Larger values of λ yield fewer features, and similarly, larger values of η yield fewer windows. When both parameters are zero, the algorithm performs no feature selection and in fact acts as a linear support vector machine classifier. The number of features yielded in this case is bounded only by the size of the training set.

4.3. Results on Simulated Data

4.3.1. Normally distributed features. This section presents the experiments used to analyze the performance of the window selection, comparing it to the original feature selection algorithm without locality and to classification without feature selection. The experiments were applied in a situation where the true distributions of the classes are known, only a number of features is relevant for classification, and the relevant features are arranged in a structured and localized manner. Two similar experiments were performed. In the first, the relevant features are arranged sequentially in a single contiguous block, and in the second, the relevant features are arranged into two disjoint contiguous blocks.

In both experiments, samples **A** and **B** were drawn from a 15-dimensional multivariate normal distribution with means μ and $-\mu$ and identity covariance matrices. The vector μ has 9 elements that are equal to zero and 6 non-zero elements. The non-zero elements are arranged into contiguous groups: in Experiment I, they form one group of 6 elements, in Experiment II they form two groups of 3 elements each:

$$\begin{aligned} \mu_{\mathrm{I}} &= \{1,1,1,1,1,1,0,0,0,0,0,0,0,0,0\}/(2\sqrt{6}) \\ \mu_{\mathrm{II}} &= \{1,1,1,0,0,0,0,0,1,1,1,0,0,0,0\}/(2\sqrt{6}) \end{aligned}$$

The denominator $2\sqrt{6}$ makes the distance between $\mu_\bullet$ and $-\mu_\bullet$ equal to 1.

For each of the means $\mu_{\mathrm{I}}, \mu_{\mathrm{II}}$, experiments were performed using training sets of sizes 30, 60, 90, and 120. For each training set size, 40 random training sets were drawn from the multivariate normal distributions $N(\mu_\bullet, \mathbf{I})$ and $N(-\mu_\bullet, \mathbf{I})$. For each training set, the feature selection and window selection algorithms were applied using different values of energy modulation parameters λ and η ($\lambda = 0.01, 0.02, \ldots, 0.2$, $\eta = 0.02, 0.04, \ldots, 0.2$). For every combination of a training set and a parameter value set, the feature selection and window selection algorithms were applied using 10 different random initializations,

and of the 10 the result with the lowest objective value was recorded. Multiple initializations were used because the fast successive linear approximation algorithm used to solve (32) does not guarantee finding the global minimum of the objective function.

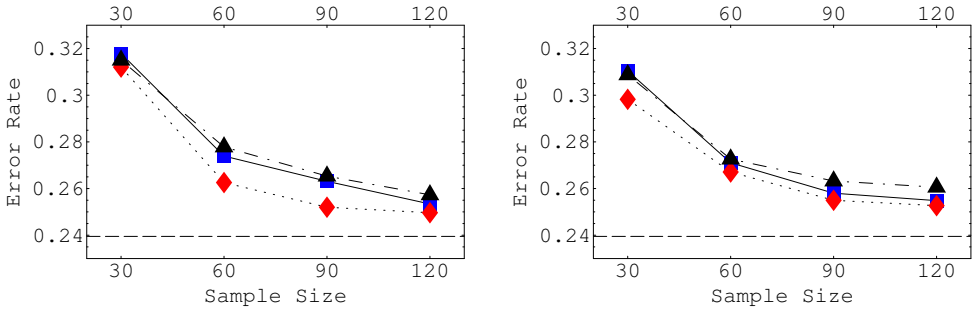
The performance of the feature and window selection algorithms on different training sets with different values of modulation parameters was measured by computing the expected generalization performance of classifiers trained on the subsets of features produced by the selection algorithms. While the feature and window selection algorithms use an L_1 support vector classifier internally, that does not dictate the choice of classifier used for subsequent discrimination. In these experiments, the Fisher linear discriminant was used for classification, because it is a maximum likelihood classifier for data sampled from a pair of multivariate normal distributions with equal covariances. In hindsight, I should have used the classifier based on the hyperplane separating the sample means, because it is the maximum likelihood classifier in the case of identity covariance matrices.

The expected generalization rate of a linear classifier is computed as follows. Let $\mathbf{w}$ be a unit length discriminant vector, and let $\{\mathbf{x} : \mathbf{w}^T\mathbf{x} = \gamma\}$ define the decision boundary of a linear classifier formed by $\mathbf{w}$. Since the underlying normal distributions are known, the expected error rate ϵ of the classifier can be computed as

$$2\epsilon = \int_{-\infty}^{\gamma} \phi(t - \mu^T\mathbf{w})\,\mathrm{d}t + \int_{\gamma}^{\infty} \phi(t + \mu^T\mathbf{w})\,\mathrm{d}t, \tag{33}$$

where $\phi(t)$ is the standard normal probability density.

Figure 4.1 shows the average expected error rate achieved by each feature selection approach for each sample size. For comparison, the error rate of a classifier based on the Fisher linear discriminant is also plotted. The window selection algorithm outperforms the original feature selection algorithm in both experiments. The difference between the generalization rates of the two algorithms is smaller in experiment II than in experiment



Experiment I. Experiment II.

FIGURE 4.1. A comparison of the average expected error rates of the window selection algorithm (red diamond, dotted line), the feature selection algorithm (blue square, dashed line), and global discriminant analysis (black triangle, solid line). The plots show the average error rates achieved with each algorithm for sample sizes ranging from 30 to 120. The results shown were computed using $\lambda = 0.03$ in both experiments, $\eta = 0.16$ in Experiment I and $\eta = 0.02$ in Experiment II.

I. As expected, the boost from window selection is stronger when the relevant features are organized into one window than when they are organized into two windows.

However, the feature selection algorithm has a lower computational complexity because its formulation as a linear programming problem has an $O(m+k+n)$ variables and $O(m^2 + k^2 + n^2)$ inequalities while the window selection algorithm generates a problem with $O(m + k + n + N)$ variables and $O(m^2 + k^2 + n^2 + n \cdot N)$ inequalities. In these experiments $N = n(n + 1)/2$, as the window set contained all the windows possible.

Figure 4.2 shows the frequency with which both algorithms find the correct feature subset in each experiment. For each sample size, it shows the number of times the feature subset was found for some parameter value, as well as the number of times that such a parameter value yielded the smallest cross-validation error. The window selection algorithm correctly selects the relevant set of features more frequently than the feature selection algorithm.

4.3.2. Synthetic shape example. The results on normally distributed data show the superiority of the window selection algorithm in a purely artificial setting. To test whether this superiority extends to shape-based classification, experiments were conducted on

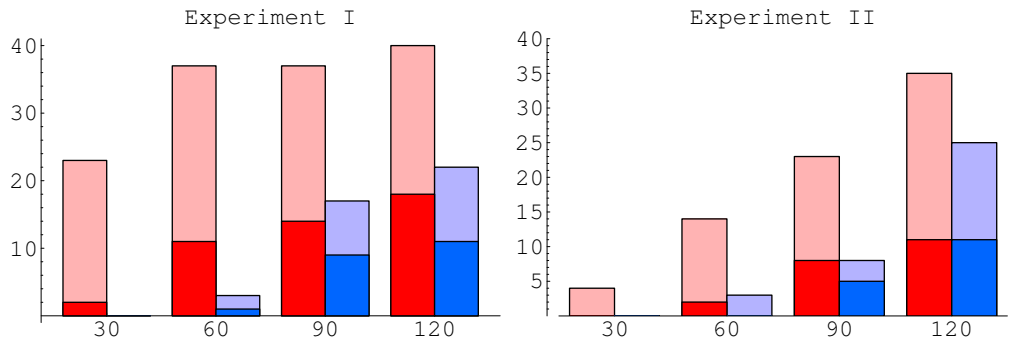


FIGURE 4.2. These charts show how many times out of 40 the two feature selection algorithms selected the correct subset of features, for a range of sample sizes. The red double bars (left) describe the window selection algorithm and the blue double bars (right) describe the feature selection algorithm. The top part of each double bar shows how many times the right feature subset was selected for *some* value of the parameters. The bottom part shows how often the right feature set was deemed optimal by cross validation.

synthetic shape data. These experiments were performed on two classes of artificial objects, which were constructed using a point distribution model of 71 corpus callosum boundaries. The point distribution model was graciously provided to us by Dr. Guido Gerig and was computed from the same corpus callosum data set as the one used in the experiments in Ch. 3.[2] The point distribution model consists of 64 boundary points that had been regularly sampled from Fourier contour segmentations of the corpus callosum. The segmentation technique [Székely et al. 1996] was used to align the objects in space and to ensure correspondence.

The two classes were generated in such a way that the differences between them are restricted to a single area of the corpus callosum. The differences between the classes were made difficult to detect by addition of a large global variability component that is common to the two classes.

The global variability component was simulated by randomly sampling members of both classes from a single Gaussian probability distribution, which was fitted to the

[2]Note that while this chapter some of the same shape data as Ch. 3, different features are used to describe the data: Ch. 3 derives features from the medial representation, and this chapter works exclusively with features derived from object boundaries.

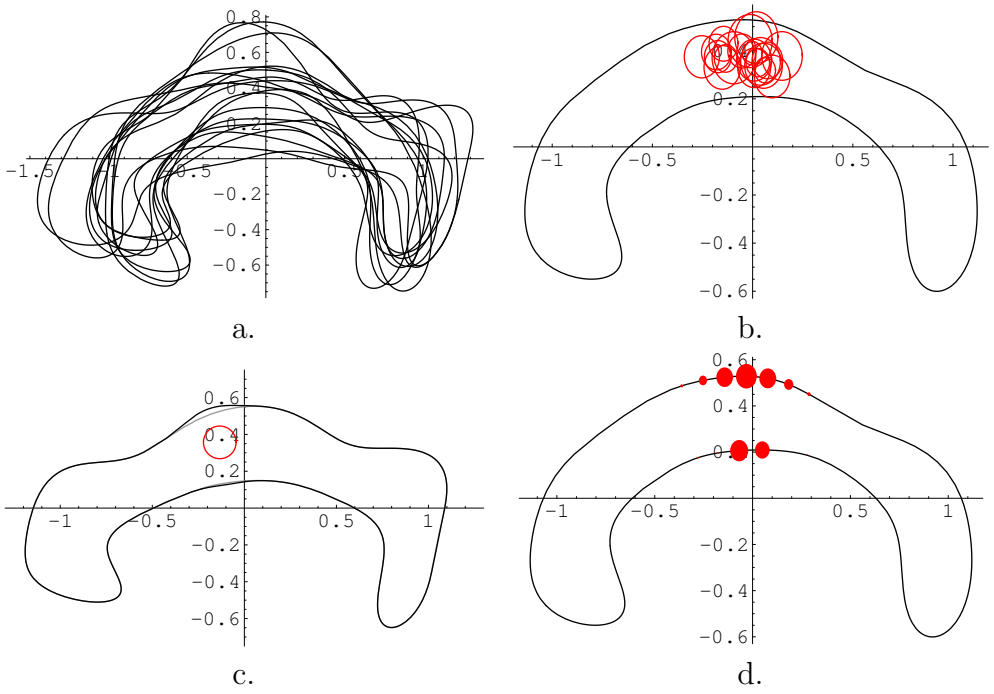


FIGURE 4.3. The generation of synthetic shapes. a. Samples from the Gaussian distribution that describes global shape variability common to both classes. b. A distribution of random tumors in relation to the average corpus callosum shape. c. The effect of a embedding a tumor into the object: the grey outline is sampled from the global shape distribution, the black outline in the deformed outline, the circle indicates the location of the warped tumor. d. The relevance measure on the features, as indicated by the relative sizes of the red dots on the boundary of the corpus callosum.

corpus callosum point distribution model using principal component analysis, following the methodology of Cootes et al. [1994]. Fig. 4.3a shows a few of the randomly generated outlines.

A local difference between the classes was induced by applying a randomized deformation to the outlines in Class II, using the following four-step procedure:

(1) A random outline O_{rnd} was sampled from the Gaussian distribution. This is the same step used to generate the outlines in Class I.

(2) A circular 'tumor', described by a point $\mathbf{c}$ and radius r, was randomly generated near a predetermined location inside of the average corpus callosum outline O_{avg}.

Fig. 4.3b shows the outline O_{avg} and the distribution of the randomly sampled tumors.

(3) A circular outward deformation field

$$\phi(\mathbf{x}) = \frac{\mathbf{x} - \mathbf{c}}{\|\mathbf{x} - \mathbf{c}\|^{\frac{3}{2}}} r \tag{34}$$

is applied to each point of O_{avg}, yielding a new outline O_{tumor}.

(4) A thin plate spline warp field, which interpolates the mapping from O_{avg} to O_{rnd}, was applied to O_{tumor}, yielding the final outline O_{def}, which then becomes a member of Class II. Fig. 4.3c shows the effect of the tumor, by superimposing the outlines O_{def} and O_{rnd}.

A single feature was computed at each of the 64 points in the point distribution model of the corpus callosum boundary. This feature was computed as the distance from the point to the origin Q. The point sets describing the corpus callosum boundaries are aligned so that the origin Q is their common center of gravity.

The approach used to produce the outlines in Class II makes it possible to measure the relative relevance of each feature in terms of how much it is affected on average by the tumor embedding . The relevance of feature i is computed as the average over all instances j in the training set of the difference between the distance to Q from the i-th point on the final outline O^j_{def} and the distance to Q from the i-th point on the intermediate outline O^j_{def}. Fig. 4.3d shows the relevance of each feature in the experiment.

The two classes were generated with 100 objects in each. Another 1000 objects were generated for testing the generalization ability of the selected feature sets. The feature selection algorithms with and without locality were applied to the classes using a broad range of modulation parameters λ and η. The L_1 support vector classifier (30) was used for measuring the generalization ability of feature subsets. Unlike the experiments in Sec. 4.3.1, there is no motivation for using the Fisher linear discriminant or the difference of

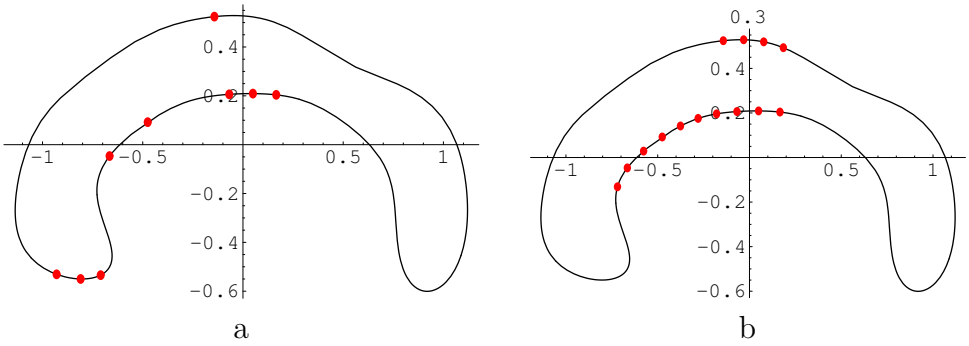


FIGURE 4.4. The best generalizing subset of features selected by the feature selection algorithm without locality (a) and by the window selection algorithm (b). The red dots on the boundary indicate selected features.

means classifier, as they are not the maximum likelihood classifiers for the shape features, which are not distributed normally.

Experiments using feature and window selection were repeated over a broad range of values of the modulation parameters λ and η. Of the resulting ranges of selected feature sets, the ones that yielded best performance on test data are reported here. Figure 4.4a shows the best feature subset for the feature selection algorithm without locality. This subset generalizes to the test data with the error rate of 0.310 and contains a large number of irrelevant features. Figure 4.4b shows the best result for window selection, which achieves the error rate of 0.300 and produces two windows, which together cover the relevant subset of features. The error rate of a classifier applied without feature selection is 0.386.

The optimal set of features yielded by window selection contains more relevant features that the optimal set reported by feature selection. Even though some irrelevant features are included in one of the two windows yielded by window selection, the results are encouraging, taking into account the small sample size, high dimensionality, and relatively large inter-population variability in the simulated data.

This experiment would have been more compelling if the training set were somehow used to determine the optimal values of the parameters λ and η. One way to do so is

to apply cross-validation to search for the best parameter values, as is done in [Bradley and Mangasarian 1998]. In my experiments I have found that this approach is highly prone to error and the parameter values that yield optimal cross-validation performance often result in suboptimal performance on test data. This is illustrated by Fig. 4.2, which shows that for the normally distributed data with one or two windows of relevant features, the parameters deemed optimal by cross-validation frequently fail to select the relevant features.

However, it is not clear that in practical applications one should be interested in finding the optimal values of λ and η. The values of these parameters directly influence the number of selected feature and the number of selected windows. The decision about the roughly desired numbers of features and windows can be made a priori, and the algorithms can be applied for a narrow range of parameters λ and η. This is the approach taken in the following section.

4.4. Results for Clinical 3D Data

4.4.1. The Hippocampus-in-Schizophrenia Data Set. The window and feature selection algorithms were applied to the study of the shape of the hippocampus in schizophrenia using exactly the data set reported in [Csernansky et al. 2002]. The data set consists of 117 subjects, 52 of whom are schizophrenia patients, and the remaining 65 are matched healthy controls. The left and right hippocampi of each subject are described using boundary meshes that consist of 6,611 vertices and 13,218 triangular faces. An example of these meshes is shown in Fig. 4.5.

These segmentations were obtained using large-deformation diffeomorphic image matching method described in [Joshi et al. 1997, Haller et al. 1997, Csernansky et al. 1998, 2002]. This method is used to establish a diffeomorphic transformation from a hippocampal template, which is based on one of the subjects in the study, to the hippocampi of the remaining subjects. The transformation adheres to the laws of fluid dynamics, and

FIGURE 4.5. An example of the hippocampus data from the schizophrenia study. Shown are, in order from left to right, top view of the left hippocampus, bottom view of the left hippocampus, top view of the right hippocampus and bottom view of the right hippocampus.

maximizes an objective function that combines terms based on optimal alignment of manually placed anatomical landmarks and the optimal match between the MR image intensities of the corresponding places in the hippocampi. Once such a transformation is established, the boundary meshes of all the hippocampi in the data set are computed by applying the transformation to the meshes extracted from the boundaries of the left and right template hippocampi. The use of anatomical landmarks ensures good correspondence between hippocampi at a global level, and the intensity match coupled with a smooth template-to-target transformation ensures good local correspondence.

Hippocampus is not a homogeneous structure but rather consists of many identifiable sub-regions, which may be affected differently by schizophrenia. Indeed, [Csernansky et al. 2002] stipulates that "the pattern of shape abnormality suggested a neuroanatomical deformity of the head of the hippocampus, which contains neurons that project to the frontal cortex". However, the statistical methodology employed in [Csernansky et al. 2002] is based on the eigenshape formulation that does not allow local specificity of shape variation. The motivation for applying feature and window selection to this data set is to find the regions of the hippocampus where the shape differences associated with schizophrenia are most significant.

4.4.2. Verification of the Results by Csernansky et al. [2002]. The data set described above was used by Csernansky et al. [2002] to show that the hippocampal shape and volume in schizophrenia patients differ from the shape and volume of hippocampi in control subjects. These results confirm earlier findings of Csernansky et al. [1998] and the findings of other researchers, such as the work of Shenton et al. [2002] on the amygdala-hippocampus complex. In order to verify that the data set was received and processed properly, I attempted to duplicate the results of Csernansky et al. The following paragraphs describe my steps, which roughly follow the procedure prescribed in [Joshi 1997].

Using the Generalized Procrustes algorithm [Gower 1975], I aligned the corresponding vertices of the hippocampal meshes, as to minimize the mean square distance to the common mean. The algorithm was restricted to shape-and-size normalization, which only translates and rotates the hippocampal meshes and retains their original sizes. The algorithm was applied separately to the left and right hippocampi in order to minimize the effect of head size.

The Cartesian coordinates of each point in both meshes were used to form features. Thus each subject was represented by $3 \times 6611 \times 2 = 39666$ features[3]. Principal component analysis was used to reduce the dimensionality of the data: the eigenvectors of the pooled covariance matrix were computed, and each feature vector was projected on the 20 eigenvectors with highest associated eigenvalues. The resulting low-dimensional data was used for classification and testing.

I began by testing how well the entire set of 20 eigenmodes can be used to discriminate between healthy subjects and schizophrenia patients. The use of a classifier based on the Fisher linear discriminant in a leave-one-out cross-validation experiment yielded the correct classification rate of 67.5%. The use of the L_1 support vector classifier (i.e., the

[3]At this step I deviated from the procedure prescribed by Joshi et al., which weights the coordinates of the vertices by the local area measure.

classifier constructed by applying the feature selection algorithm with λ equal to zero) yielded the correct classification rate of 62.4%.

Csernansky et al. use a logistic regression model to find a subset of the first 20 eigenmodes that can be used to optimally separate the two classes. They do so in a robust setting, using 10-fold cross-validation [Stone 1974]. The data set is divided into 10 roughly equal parts, and in a series of 10 experiments, one of the parts is set aside while linear logistical regression is performed on the remaining subjects, yielding a small number of selected eigenmodes. A classifier restricted to those eigenmodes is constructed and applied to the withheld data and the error rate of the classifier is recorded. The average error rate over all 10 experiments is then reported. Csernansky et al. report the average correct classification rate of 68.3% and note that the eigenvector sets selected in each experiment are similar.

I replicated this experiment by applying feature selection instead of the logistic regression model to the 20 eigenmode features inside of a 10-fold cross-validation experiment. However, in order to gauge the significance of this cross-validation technique, I repeated the cross-validation 10 times, each time using a different random partition of the data set. The resulting average correct classification rates ranged between 59.0% and 70.9%. The large value of dispersion indicates that the significance of a single 10-fold cross-validation experiment is low. Using feature selection inside of a more expensive leave-one-out experiment, I obtained the correct classification rate of 73.5%. In both types of experiments, the value of the parameter λ from (24) was chosen to return roughly 4 features on average. The features returned in each application of feature selection were highly similar, just as in the experiments performed by Csernansky et al. In fact, in 113 of the 117 leave-one-out experiments, the selected eigenmodes were $1, 7, 16$, and 20.

Csernansky et al. also apply the logistic regression model to the entire data set and report the maximum between-group discrimination rate of 76% for the resulting set of eigenmodes. My version of this experiment using feature selection yielded the similar

leave-one-out correct classification rate of 77.3%. However, the significance of these numbers is questionable because the eigenmodes were selected using the knowledge of class membership of each subject in the data set may not generalize.

4.4.3. Reducing Mesh Dimensionality for Window Selection. In order to use window and feature selection to produce regions large enough to cover 10%-20% of the hippocampal surface, the number of features was reduced from nearly $40,000$ that result from using the x, y, z coordinates of each mesh vertex as features, to 160 localized summary features, which describe small patches on the surface of the hippocampus. The reduction was necessary because window selection and feature selection algorithms yield fewer features than there are subjects in the training set (as mentioned in Sec. 4.2.3). Such a limit on the number of selected features among the initial features would not allow even 1% of the hippocampal surface to be covered with selected regions. Moreover, the computational cost of applying window selection to nearly $40,000$ features would be prohibitive.

Patch features were computed as follows. The sets of 117 left and 117 right meshes were aligned using the Generalized Procrustes algorithm [Gower 1975] restricted to translation and orientation. In the process, the mean left and right hippocampal meshes were computed. Each mesh was subdivided into 80 patches of roughly equal area using the *METIS* graph partitioning software [Karypis and Kumar 1998] on a graph whose vertices correspond to the mesh triangles and are weighted by the average areas of the triangles. The partitioned left and right mean meshes are shown in Fig. 4.6.

Each patch was represented with a single summary feature, which measures the average inward or outward deformation of the patch with respect to the mean mesh. The summary feature computed for the j-th patch in the i-th subject is given by

$$f_{ij} = \frac{\sum_{k \in \mathcal{P}(j)} (\mathbf{x}_{ik} - \bar{\mathbf{x}}_k)^T \, \mathbf{N}_k \, \Delta A_k}{\sum_{k \in \mathcal{P}(j)} \Delta A_k}, \tag{35}$$

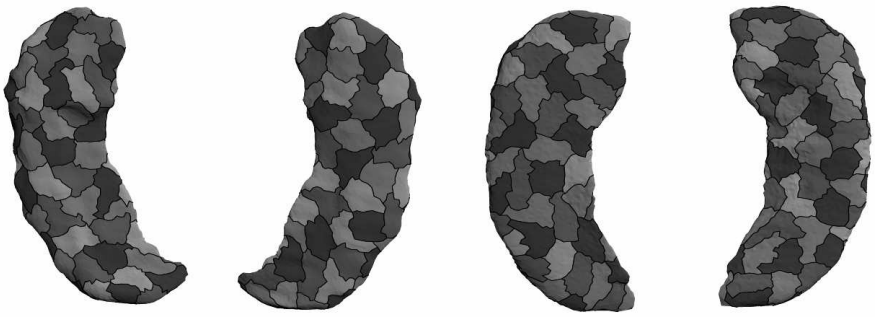


FIGURE 4.6. Mean left and mean right hippocampal meshes partitioned into 80 patches each. The meshes are shown from superior and anterior viewpoints.

where $\mathcal{P}(j)$ is a set of indices of the vertices belonging to the j-th patch, $\bar{\mathbf{x}}_k$ and $\mathbf{N}_k$ are the the position and approximate unit normal of the k-th vertex in the mean shape, and ΔA_k is the area element, computed as one third of the combined area of all triangles adjacent to the k-th vertex in the mean shape. The numerator in 35 is an approximation of the surface integral of the normal displacement from the mean shape, and the denominator is just the area of the patch on the mean shape.

In order to gauge how much information was lost by replacing nearly 40,000 features with only 160 summary features, the latter were analyzed using the same PCA-based approach that was used in section 4.4.2: the features were projected onto the first 20 eigenmodes and subjected to classification and cross-validation. The leave-one-out rate for the L_1 support vector classifier applied to the entire set of 20 eigenmodes yielded a 67.5% success rate, while Fisher linear discriminant yielded 66.7%. When feature selection was used in the leave-one-out loop to select the few best eigenmodes, the success rate of 69.2% was obtained. The sets of eigenmodes selected in each iteration of leave-one-out cross-validation were homogeneous, with eigenmodes $1, 12, 13$, and 15 being selected in 111 of the 117 experiments.

4.4.4. Results of Feature Selection and Window Selection. An alphabet of windows was defined over the patch summary features using the transitive distance function,

which counts the number of patch edges that separate any two patches. Under this function, single patches form windows of size 0 and sets of mutually adjacent patches (typically consisting of three patches) form windows of size 1. For computational efficiency, windows of larger size were not included in the alphabet.

Feature selection and window selection algorithms were applied to patch summary features in a series of leave-one-out cross-validation experiments. In each leave-one-out iteration, one subject was removed from the data set, the selection algorithm was applied to the remaining subjects, an L_1 support vector classifier was constructed in the subspace spanned by the selected features, the left out subject was assigned a class label by the classifier, and this class label was compared to the true class label of the left out subject. The average correct classification over 117 leave-one-out iterations was recoded. The feature selection and window selection experiments were repeated for different values of modulation parameters λ and η. Table 4.1 shows the results of these experiments.

Recall that in [Csernansky et al. 2002], using a 10-fold cross-validation methodology, a similar classification rate of 68.4% is reported. The methods in [Csernansky et al. 2002] are based on eigenanalysis of the entire set of $40,000$ features. The results in Table 4.1 show that with intelligent feature selection a similar classification rate can be achieved with only 160 summary features. The feature selection methodology also specifies the local regions of the hippocampus that are significant for discrimination.

The top row of Fig. 4.7 shows the ten patches that were selected most frequently in the 117 leave-one-out experiments conducted with the feature selection algorithm with $\lambda = 0.16$. The second row of Fig. 4.7 shows the ten most frequently selected patch windows in the window selection experiment with $\lambda = 0.12$ and $\eta = 0.08$. Window selection results in fewer isolated features than feature selection. For reference, the bottom row of Fig. 4.7 plots the p-values of mean difference hypothesis tests computed at each patch. No correction for the repeated nature of tests has been applied. While the pattern of patches selected by the window and feature selection algorithms closely resembles the

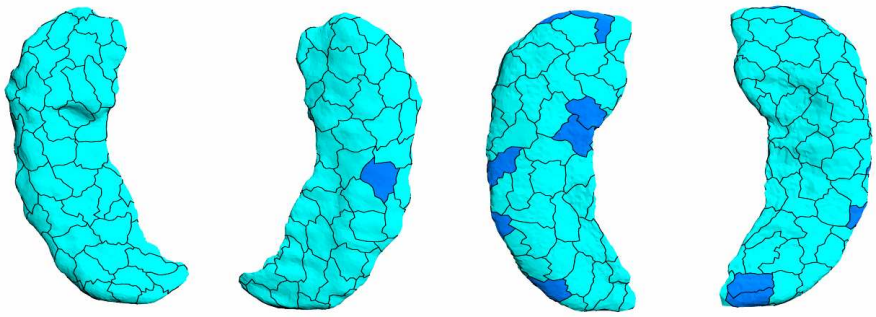
TABLE 4.1. Cross-validation results for clinical data. Each row summarizes a leave-one-out experiment performed on patch summary features using a different pair of modulation parameters η and λ. For each experiment, the average number of selected windows (except for $\eta = 0$), the average number of selected features, and the average leave-one-out correct classification rate are shown.

η	λ	Avg. Win.	Avg. Feat.	Corr. Rate (%)
0.0	0.04		22.9	55.6
0.0	0.08		16.4	65.0
0.0	0.12		7.5	65.0
0.0	0.16		4.6	68.4
0.04	0.04	11.8	28.7	68.4
0.04	0.08	8.5	19.3	69.2
0.04	0.12	4.2	8.5	62.4
0.04	0.16	2.1	4.0	54.7
0.08	0.04	10.8	28.1	61.5
0.08	0.08	5.7	13.5	64.1
0.08	0.12	2.8	6.1	64.1
0.08	0.16	1.6	2.9	59.0
0.12	0.04	9.2	24.5	68.4
0.12	0.08	3.9	9.9	62.4
0.12	0.12	2.1	4.7	57.3
0.12	0.16	1.4	2.8	61.4

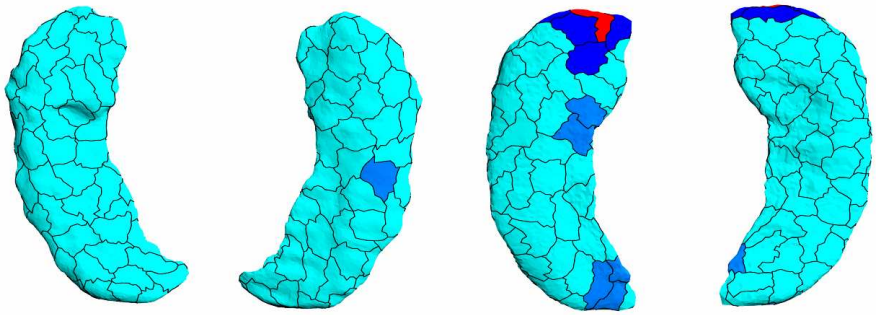
pattern of patches with low p-values, the selected patches do not correspond to the patches with lowest p-values. As indicated in [Csernansky et al. 2002], the head of the right hippocampus was shown by window selection to be most relevant for discrimination.

4.5. Discussion and Future Work

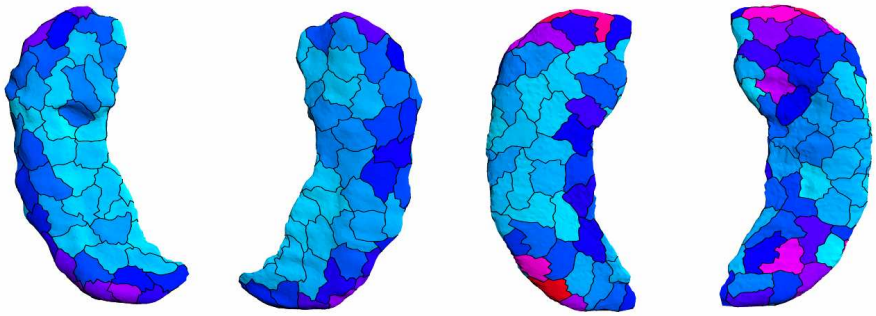
Let us summarize the results obtained in this chapter. In three different kinds of classification experiments, the results show some aspect of superiority of the window selection method over the generic feature selection method. In the first set of synthetic experiments, which have normally distributed features with relevant features chosen a priori in a localized pattern, the window selection method performed as expected, yielding better generalization rates and correctly selecting the relevant feature subsets more often than the feature selection algorithm. This result is not theoretically surprising, but it is



Patches selected most frequently by feature selection



Patches selected most frequently by window selection



Patch wise p-values

FIGURE 4.7. Top row: ten patches that were selected most frequently during leave-one-out validation of feature selection. The meshes are shown from superior and anterior viewpoints. Second row: ten windows that were selected most frequently during leave-one-out validation of feature selection (some of the windows overlap, and patches that belong to more than one window are shaded darker on the cyan-red hue scale). Bottom row: p-values of the mean difference tests computed at each patch; the negative logarithm of the p-values is displayed using the cyan-red hue scale (cyan = no significance, red = high significance).

rather a way of verifying that window selection really does work in a situation when it should.

The second set of synthetic experiments involved simulated corpus callosum data and was designed to test whether or not the advantages of using window selection on localized features extends to shape data. In these classification experiments the set of relevant features was constructed by embedding a random bump at the midbody of the corpus callosum in one of the two classes, causing the relevant features to be localized to two 'fuzzy' windows. The classification task in these experiments was made rather difficult by making the variability common to the classes be much more significant than the inter-class differences caused by the bump embedding. Although neither selection algorithm did a perfect job of identifying the relevant subset of features, the window selection algorithm showed a clear superiority over the feature selection algorithm, returning a pair of feature windows that overlapped with the true 'fuzzy' windows, while the feature selection algorithm returned a large number of irrelevant features located far away from the corpus callosum midbody. The superiority of the window selection algorithm in terms of generalization of the classifier constructed on the selected feature set to test data was marginal when compared with feature selection and significant when compared to classification on the whole set of features. Again, these results are not surprising, but affirming. They also suggest that the advantage of using window selection may lie more in the area of discovering the locations with significant inter-class differences rather than improving the generalization ability of classifiers.

The third set of experiments, performed on clinical hippocampus data is of most interest. The truth, i.e., the relevant set of features, is unknown in this experiment, even though some indications of the significance of the head of the right hippocampus have been made in the literature [Csernansky et al. 2002]. The results of applying window and feature selection algorithms to patch summary features did not show significant improvement in terms of classifier generalization rate over the global PCA-based classification. In

fact, the best leave-one-out cross-validation rates obtained by window selection, feature selection, and feature selection on eigenmodes lie in the same ballpark of 69%. However, in terms of selecting the relevant feature sets, the results on patched hippocampus data are very encouraging. For example, in the experiment shown in Fig. 4.7, the windows selected most often by the window selection algorithm are nicely localized to three clusters of patches on the right hippocampus and a single patch on the left hippocampus, with a clear advantage of the head of the right hippocampus. In contrast, the feature selection results are more scattered around the hippocampi and do not show potential areas of relevance as strongly.

It is unlikely that a classification technique will one day make it possible to accurately diagnose schizophrenia only on the basis of hippocampal shape. Therefore, the goal in developing the window selection algorithm was not so much to build a better classifier but rather to find the regions of the hippocampus that are significant for discrimination. With respect to this goal, the results presented in this paper are encouraging. However, these results require further validation using a different hippocampal data set. I plan to perform this validation in the future.

I also plan to perform window and feature selection on hippocampal patches selected manually on the basis of biological homogeneity and function. The use of selection algorithms on anatomically significant patches could open new insights into schizophrenia.

On the theoretical front, I plan to extend this paper's framework to select features in a hierarchical manner. Selected patches would be further partitioned into smaller patches, and the selection algorithms would be performed again on the residuals, resulting in a high-resolution set of selected features. Hierarchical feature selection would eliminate the information loss incurred by reduction to patch summary features.

In order for window selection to be integrated into the medially based coarse-to-fine shape characterization framework developed in Ch. 3, the nuances of applying selection to features derived from the medial representation need to be explored. The use of

geometrically descriptive medial features forces one to face the problem of having more than one feature describe each location in the object. This problem can be tackled simply by organizing a window alphabet in a way that disallows partial selection of the features that describe a single location. However, it is also conceivable to define non-zero inter-feature distances between such features. Moreover, the combined use of medial features that describe the same locations at different levels of detail (coarse-level and residual fine-level features) requires considerable theoretical consideration. Some thoughts in this direction are presented in Ch. 6, which describes a conceptual framework that would apply feature selection in a coarse-to-fine manner to features derived from the continuous medial representation, which is described in the next chapter.

CHAPTER 5

Continuous M-Reps for Geometric Object Modeling

> Must a name mean something? Alice asked doubtfully. Of course it must, Humpty Dumpty said with a short laugh: my name means the shape I am - and a good handsome shape it is, too. With a name like yours, you might be any shape, almost.
>
> Lewis Carroll

This chapter is based on a paper [Yushkevich et al. 2003] that was written together with P. Thomas Fletcher, Dr. Sarang Joshi, Andrew Thall, and Dr. Stephen M. Pizer. The coauthors' contributions were invaluable for the conception, implementation, and publication of the research described below. Neither would this work have been possible without the theory [Damon 2002] and extensive help of Dr. James Damon.

This chapter describes a novel continuous medial representation for object geometry and a deformable templates method for fitting the representation to images. The representation simultaneously describes the boundary and medial loci of geometrical objects, always maintaining Blum's symmetric axis transform (SAT) relationship. Cubic b-splines define the continuous medial locus and the associated thickness field, which in turn generate the object boundary. The chapter describes geometrical properties of the representation and presents a set of constraints on the b-spline parameters. The 2D representation encompasses branching medial loci; the 3D version can model objects with a single medial surface, and the extension to branching medial surfaces is a subject of future research. The representation and the associated segmentation procedure were

used to segmenting 2D and 3D medical images, and the results are presented in this chapter.

5.1. Introduction

Medial loci, or skeletons, have enjoyed wide use in computer vision and medical image analysis because they provide important intuition about shape and formation of biological and anatomical objects. Medial loci naturally divide objects into a hierarchy of simple figures and describe the inherent symmetry and local thickness of each figure.

As noted in Sec.2.1.4, medial loci of objects have traditionally been computed from discrete boundary-based descriptions by skeletonization algorithms. Such boundary descriptions yield medial loci with a complex branching structure. For instance, a skeleton constructed using Voronoi diagrams has roughly the same number of branches as there are vertices in the discrete boundary description. Methods that simplify and regularize skeletons can eliminate unstable branches and yield object-relevant medial loci [Kimia et al. 1995, Siddiqi et al. 1999b, Ogniewicz and Kübler 1995, Näf et al. 1996]. Nevertheless, the boundary-to-medial transformation is inherently unstable; the resulting branching topology is sensitive to slight boundary perturbations, especially at the regions known as ligatures [Blum and Nagel 1978, August et al. 1999].

The continuous m-reps approach described in this chapter, as well as the general m-reps approach, fall into a category of *synthetic* medial representations that use the medial locus itself as a model from which geometric properties of an object, such as its boundary, are derived. Synthetic medial representations describe the way in which the medial locus of an object is decomposed into branches (i.e., medial branching topology) and define each branch of the medial locus using a few parameters. A smooth parameterized thickness field is defined over the entire medial locus.

The medial locus and the associated thickness field synthesize a stable object boundary by inverting the skeletonization process. The generated boundary is equivalent to

the envelope of spheres (or disks) placed at each point in the medial locus with the radius prescribed by the associated thickness value. Synthetic medial representations establish a correspondence between each point on the medial locus and a pair of points on the generated boundary (i.e., the boundary pre-images). Synthetic medial representations enforce a fixed medial branching topology and provide a *simultaneous* description of the medial locus and the object boundary.

The continuous m-rep approach is a synthetic medial representation that uses cubic b-splines to model both the medial manifolds and the associated thickness field. The approach relies on constraints developed by Damon [2002] to guarantee that the generated boundary surface is a closed, connected, non-singular manifold with curvature continuity. The current state of the method allows representation of 2D objects with branching medial topology and 3D objects with a single medial surface. The extension to branching medial surfaces is the subject of ongoing research.

Continuous m-reps essentially are a continuous extension of discrete *m-reps* described in Sec. 2.1.7. The results presented in this chapter show that continuous m-reps can be applied to segment anatomical objects in medical images using the same deformable templates framework that is used by Joshi et al. [2002] for discrete m-rep segmentation. These results include automatic segmentation of a vertebral image from a CT slice using a 2D model with branching medial topology and in 3D, automatic deformation of a template model of the hippocampus to fit manually segmented magnetic resonance images of the human brain.

Continuous m-reps were developed with the ultimate goal of improving present medially based methods for statistical shape analysis. Methods that use discrete m-reps with a fixed branching topology to describe a population of objects [Styner and Gerig 2001a, Yushkevich et al. 2001, Fletcher et al. 2003a] estimate probability distributions of statistical features derived from medial atoms. These distributions are used to generate new instances of m-reps, to visualize the primary modes of variability in the population

in terms of bending or thickening of objects, to perform classification on the basis of shape and to pinpoint locations in objects where shape variability is most pronounced. Continuous m-reps augment shape analysis methods by allowing arbitrary sampling of medial loci. The continuous medial representation makes it possible to elastically model and optimize correspondences between features of different objects in the population.

The details of the continuous m-rep approach are described in the following five sections. Section 5.2 summarizes elements of related research in boundary and medial object representation. Section 5.3 develops the geometric foundation for continuous medial modeling and defines the generative medial b-spline model. Section 5.4 describes the procedure that computes the optimal parameters of the model for use in automatic segmentation. Section 5.5 presents the preliminary segmentation results on real medical images. Section 5.6 describes the present limitations of the method, and compares the b-spline medial representation with discrete m-reps.

5.2. Background

Local boundary representations, polygonal or higher order approximations of the boundary surface, are the principal shape descriptors in computer graphics, image analysis and shape characterization. These representations by themselves provide no *a priori* global information, such as the figural hierarchy of an object, and thus do not provide an intuitive framework for shape analysis and figure-based deformation. Nevertheless local primitives are used widely, for instance for elastic deformable modeling of 3D objects of arbitrary topology [Delingette 1994].

Global boundary shape descriptions have found use in image analysis via deformable models and in shape analysis. Staib and Duncan [1992] represent boundaries of three-dimensional objects as weighted sums of Fourier components. Székely et al. [1996] use a spherical harmonics decomposition to represent three-dimensional objects of spherical topology. Carr et al. [2001] use radial basis functions to represent boundaries. Styner

and Gerig [2001a] combine the spherical harmonic SPHARM representation with discrete m-reps and computes statistics on the combined representation, yielding a comprehensive statistical analysis of shape.

Object description via skeletons, i.e. medial representation, has been used increasingly in computer graphics and CAD. Bloomenthal and Shoemake [1991] and Sherstyuk [1999] develop implicit surfaces based on convolution over medial skeletons. Markosian et al. [1999] apply *skinning* of implicit fields around polyhedral skeletons. Gascuel [1993] develop a system for animation and collision detection based on rigid articulated skeletons fleshed by spline-based deformable boundary surfaces. Storti et al. [1997] and Blanding et al. [1999] use a skeleton-based object representation for CAD style applications: 3D geometric model synthesis, generation of boundary surfaces at varying levels of detail, and morphing. The Teddy system [Igarashi et al. 1999] uses the medial spines to drive intuitive shape modeling based on hand sketching. Chen et al. [1999] use multiscale medial models based on a sampled skeleton for guiding volume rendering. Thall et al. [2000] propose discrete m-reps as a general geometric modeling primitive for 3D design in computer graphics and CAD, using displaced subdivision surfaces for fine-scale boundary description.

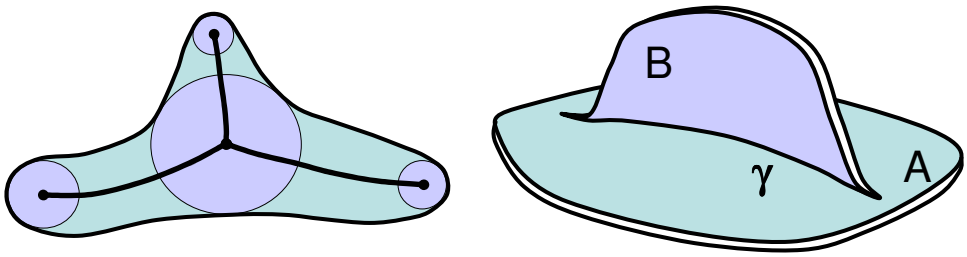
Use of medial representations for deformable modeling in computer vision was pioneered by work of Tsao and Fu [1984], where discrete skeletons computed by distance transform are stochastically manipulated and a discrete boundary is regenerated. Zhu and Yuille [1996] developed the comprehensive FORMS system that automatically divides 2D objects into simple parts and represents these parts medially, incorporating statistical shape information. The concept of discrete m-reps and their use in medical image analysis are discussed extensively in Sec. 2.1.7. The difference between the deformable approaches mentioned here and the continuous m-reps approach is that the latter enforces Blum's symmetric axis transform (SAT) relationship between a continuous deformable medial model and the boundary generated by it.

Differential geometric properties of medial axis transforms, boundary generation from continuous skeletons, and associated validity constraints have been studied extensively in the literature. In particular, Nackman [1982], Vermeer [1994], and Gelston and Dutta [1995] explored relationships between the 3D SAT and the curvature of its implied boundary. Hoffmann and Vermeer [1996] and Teixeira [1998] described validity conditions that continuous medial surfaces must satisfy in order to exist and to imply non-intersecting boundaries. Giblin and Kimia [2000] studied local differential geometry of various kinds of points forming 3D symmetry sets. Damon [2002] studied the geometry of offset surfaces generated by a multi-valued vector field defined on a set of connected manifolds; the results presented in the following section represent a special case of Damon's theory.

5.3. Method

This section presents the details of continuous m-reps. Subsection 5.3.1 describes the differential geometry of continuous medial manifolds. Subsection 5.3.2 presents the constraints that must be satisfied by a continuous medial model in order for the resulting boundary to form a closed non-singular manifold. In subsection 5.3.3 cubic b-splines are used to implement the continuous medial representation, incorporating the constraints.

The geometric notions presented in the first two subsections constitute a special case of the results recently developed by Damon. Damon [2002] studies a class of manifolds formed by offsetting a branching medial locus by a multi-valued vector field. He describes the differential geometry of the offset manifold in terms of the shape operator of the vector field and gives a set of necessary and sufficient constraints that both the medial surface and the vector field must satisfy in order to produce a continuous boundary without singularities.



(a) (b)

FIGURE 5.1. Branching medial loci in in 2D and 3D. (a) In 2D, three branches generically join at shared endpoints. Approaching the shared endpoint, the thickness (radius) values associated with each branch converge to a common value. (b) In this 3D example, two medial surfaces are connected along a seam γ. The seam forms a crease in A and is part of the edge of B.

5.3.1. Differential Geometry of Medial Manifolds. A synthetic continuous medial representation (*cm-rep*) defines the medial locus of an object as a set of connected parameterized manifolds called *medial manifolds*. A thickness value is associated with each point of each medial manifold.

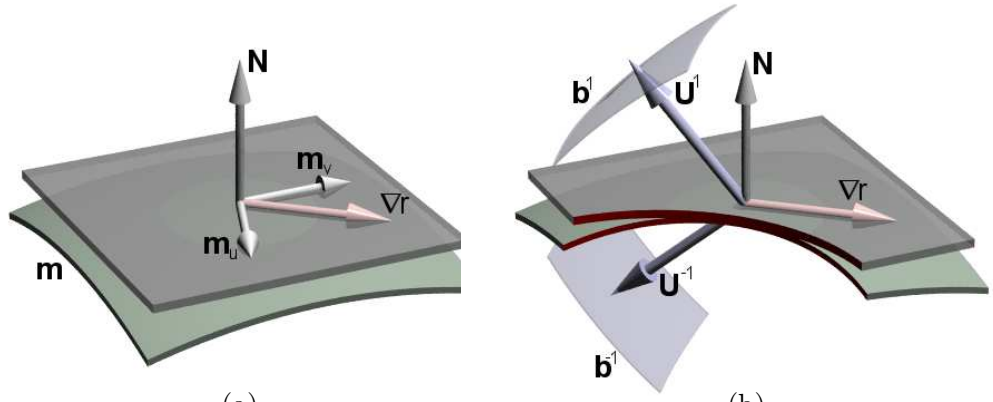
Medial manifolds in 2D are smooth curve segments connected to each other at endpoints. Medial curve segments connect when three of them come together at a shared endpoint, as demonstrated in Fig. 5.1a. The thickness value is equal at the shared endpoint for all three branches. A joining of more than three branches at a point is non-generic.

In general, a 3D medial locus is a set of connected medial surfaces and space curves. Tubular objects whose medial loci are curves are non-generic and are not dealt with in this dissertation. Each medial surface is bounded by a closed curve that is called the *edge*. In particular, the part of an edge that is not part of any other medial surface is called the *free edge*.

Medial surfaces connect along shared curves that are called *seams*. Seams either form a part of the edge of a medial surface or form a crease (a discontinuity in the surface

TABLE 5.1. Medial notation.

Symbol	Description
$\mathbf{m}$	Medial surface
r	Radial scalar field
u, v	Parametrization of $(\mathbf{m}, r)$
$\mathbf{b}^t$	Boundary counterparts of $(\mathbf{m}, r)$
t	Indexes the two parts $(-1, 1)$ of the implied boundary.
$\mathbf{u}^t$	Unit normal to the boundary, also the direction from a point on $\mathbf{m}$ to its boundary counterpart.
$\mathbf{n}$	Unit normal to the medial surface.



(a) (b)

FIGURE 5.2. Elements of 3D medial geometry. (a) The medial surface $\mathbf{m}$ and its tangent plane at a point; the non-orthogonal frame $(\mathbf{m}_u, \mathbf{m}_v, \mathbf{n})$, and vector ∇r, the gradient of the thickness scalar field. (b) Implied boundary surfaces $\mathbf{b}^1, \mathbf{b}^{-1}$ and the vectors $\mathbf{u}^1, \mathbf{u}^{-1}$ that point from the medial surface to the boundary and are normal to the boundary.

normal) on the medial surface. Medial surfaces are smooth except at seams. Fig. 5.1b, shows the seam γ that is a part of the edge of A and a crease on B. Geometric aspects pertaining to 3D cm-reps with connected medial surfaces are discussed below but the spline-based implementation is limited to models with a single smooth medial surface.

Formally, let $\mathcal{O}$ be some geometric object with a closed continuous boundary. The cm-rep of $\mathcal{O}$ is an approximation of its medial axis. The cm-rep of $\mathcal{O}$ is considered *valid* if it forms an exact medial axis of a geometric object that has the same topology as $\mathcal{O}$, and whose boundary is closed, continuous in curvature, and non-singular. This boundary

is called the *implied boundary* of a valid cm-rep. The accuracy with which a cm-rep describes $\mathcal{O}$ is measured in terms of differences between $\mathcal{O}$ and the object formed by the implied boundary of the cm-rep.

In n-dimensional space, the medial locus is described by a set of control parameters that define C^2 functions $(\mathbf{m}, r) : D \rightarrow \mathbb{R}^n \times \mathbb{R}^+$ on a closed domain $D \subset \mathbb{R}^{n-1}$. The medial manifold is defined by the spacial component $\mathbf{m}$, and r defines the thickness field on $\mathbf{m}$.

The boundary generated by a cm-rep is constructed by inverting the SAT. Spheres (or disks) of radius $r(u, v)$ are placed at each location $\mathbf{m}(u, v)$ on the medial manifold. The generated boundary is the envelope of such a family of spheres or disks. In 3D, the points $\mathbf{x} \in \mathbb{R}^3$ that belong to this two-parameter family of spheres are defined by the implicit equation

$$f(\mathbf{x}, u, v) = |\mathbf{x} - \mathbf{m}(u, v)|^2 - r(u, v)^2 = 0 \ . \tag{36}$$

At the points on the envelope f must satisfy

$$f = 0 \ , \quad f_u = 0 \ , \quad f_v = 0 \ . \tag{37}$$

This system of equations yields a definition of the boundary that assigns a pair of boundary positions to each (u, v). Each sphere in the family generically touches the envelope at two points on opposite sides of the medial axis (with the exception special points where the two sides of the boundary come together, which will be discussed later). The points of tangency are called the *boundary counterparts* of a point on the medial axis. The two boundary counterparts are indexed by $t \in \{-1, 1\}$, are denoted as $\mathbf{b}^t(u, v)$, and are

expressed as:

$$\mathbf{b}^t = \mathbf{m} + r\mathbf{u}^t \ , \tag{38}$$

$$\mathbf{u}^t = -\nabla r + t\sqrt{1 - \|\nabla r\|^2}\mathbf{n} \ , \tag{39}$$

where $\mathbf{n}$ is the unit normal on the medial surface and $\mathbf{u}^t$ is the unit normal on the boundary surface. The vector ∇r is the gradient of the thickness scalar field on the medial surface:

$$\nabla r = \begin{bmatrix} \mathbf{m}_u & \mathbf{m}_v \end{bmatrix} \mathbf{I}_\mathbf{m}^{-1} \begin{bmatrix} r_u \\ r_v \end{bmatrix} \ , \tag{40}$$

where $\mathbf{I}_\mathbf{m}$ is the metric tensor on the medial surface.

The projections of both boundary counterparts onto the medial tangent plane lie in the negative ∇r direction. The distance from each counterpart to the medial tangent plane is $r\sqrt{1 - \|\nabla r\|^2}$; hence the boundary counterparts of $\mathbf{m}$ are defined only if

$$\|\nabla r\| \leq 1 \ . \tag{41}$$

This becomes the first constraint on the radial field.

The vector $\mathbf{u}^t$ in the direction from a medial point to its boundary counterpart is normal to the implied boundary. This ensures that the medial surface is the SAT of its implied boundary. This property is used to express the curvature tensor of the implied boundary in terms of second derivatives of $(\mathbf{m}, r)$.

The second derivative tensor of the implied boundary surface is

$$\mathbf{II}_{\mathbf{b}^t} = \begin{bmatrix} \mathbf{b}^t_{uu} \cdot \mathbf{u}^t & \mathbf{b}^t_{uv} \cdot \mathbf{u}^t \\ \mathbf{b}^t_{vu} \cdot \mathbf{u}^t & \mathbf{b}^t_{vv} \cdot \mathbf{u}^t \end{bmatrix} = - \begin{bmatrix} \mathbf{b}^t_u \cdot \mathbf{u}^t_u & \mathbf{b}^t_v \cdot \mathbf{u}^t_u \\ \mathbf{b}^t_u \cdot \mathbf{u}^t_v & \mathbf{b}^t_v \cdot \mathbf{u}^t_v \end{bmatrix} . \tag{42}$$

The metric tensor of the implied boundary is

$$\mathbf{I}_{\mathbf{b}^t} = \begin{bmatrix} \mathbf{b}_u^t \cdot \mathbf{b}_u^t & \mathbf{b}_u^t \cdot \mathbf{b}_v^t \\ \mathbf{b}_v^t \cdot \mathbf{b}_u^t & \mathbf{b}_v^t \cdot \mathbf{b}_v^t \end{bmatrix} . \tag{43}$$

The principal curvatures and principal directions of the implied boundary, which are the eigenvalues and eigenvectors of $\mathbf{II}_{\mathbf{b}^t}\mathbf{I}_{\mathbf{b}^t}^{-1}$, are expressed in terms of first derivatives of $\mathbf{b}^t$ and $\mathbf{u}^t$, and following (38), in terms of second derivatives of $\mathbf{m}$ and r.

At a point $\mathbf{m}$ on the medial surface where $\|\nabla r\| = 1$ the component of $\mathbf{u}^t$ in the direction normal to $\mathbf{m}$ is 0, following (39). The two boundary counterparts collapse to a common point on the tangent plane of the medial surface. The square root in (39) forces the derivative of $\mathbf{u}^t$ to asymptote at such points, and the curvature tensor on the surface can not be computed in terms of second derivatives of $\mathbf{m}$ and r.

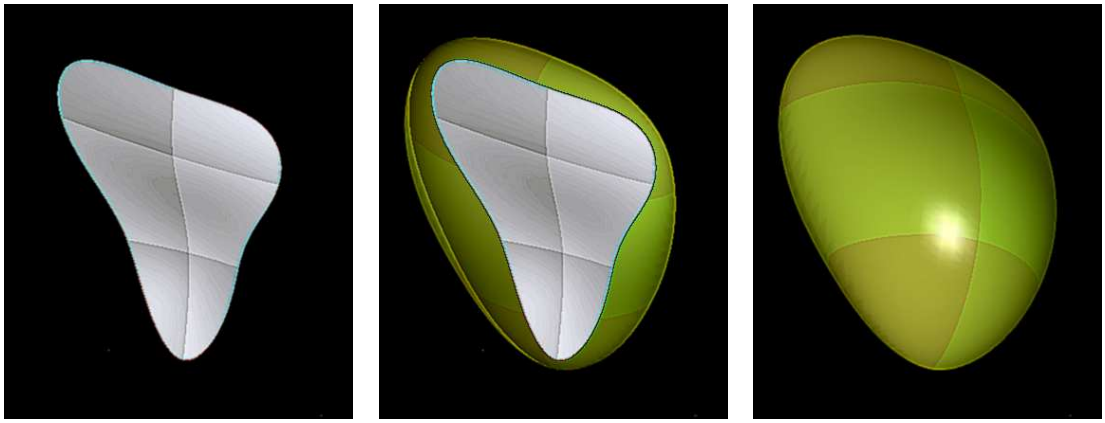
The definition of the implied boundary in 2D is analogous to 3D. Equations (38), (39) do not change, and the vector ∇r is given by

$$\nabla r = \frac{dr}{ds}\mathbf{t} = \frac{r'}{\sqrt{x'^2 + y'^2}}\mathbf{t} \ , \tag{44}$$

where s is the arclength of the medial curve, $\mathbf{t}$ is the unit tangent vector on the medial curve, and the primes are derivatives taken with respect to the parameter u.

5.3.2. Constraints on Medial Manifolds. The functions $(\mathbf{m}, r)$ must satisfy constraints that fall into two categories. The first category ensures that the cm-rep generates a closed connected boundary. The second category ensures that the boundary is non-singular. This section focuses on 3D constraints and the simplified 2D case is mentioned in the passing.

Each medial surface comprising a cm-rep generates two boundary halves, $\mathbf{b}^1$ and $\mathbf{b}^{-1}$. The constraints that ensure that the boundary halves to meet forming a closed surface are expressed differently at free edges, smooth interior points, seams and at the ends



(a) (b) (c)

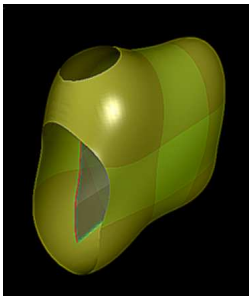
FIGURE 5.3. Examples of 3D cm-reps. (a) A medial surface. (b) Same medial surface and its implied boundary half $\mathbf{b}^1$. (c) Same medial surface with both boundary halves $\mathbf{b}^1, \mathbf{b}^{-1}$ that form a closed surface.

of seams. The spheres placed at these different classes of points have distinct orders of contact with the implied boundary, and have different geometric properties [Giblin and Kimia 2000].

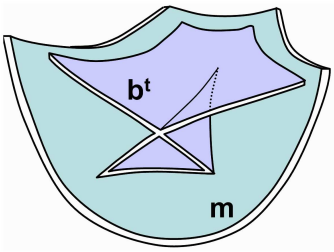
At free edges the boundary halves implied by the same medial surface connect. Recall from section 5.3.1 that the boundary counterparts of a point on a medial surface coincide when $\|\nabla r\| = 1$. Thus this condition must hold at points on the free edge. Fig. 5.4a shows an example of a medial surface that violates the free edge constraint.

At smooth points interior to a medial surface the two boundary counterparts are disjoint; otherwise the implied boundary does not form a closed manifold. Following (41), the constraint at these points is the strict inequality $\|\nabla r\| < 1$.

At a seam point three smooth subsurfaces $\mathbf{m}_i, i \in \{1, 2, 3\}$ come together at an angle to each other. For example, in Fig. 5.1b the opposite sides of the crease in A and the surface B form the three smooth subsurfaces. The six boundary halves implied by the three subsurfaces connect in such a way that $\mathbf{b}_i^1$ meets $\mathbf{b}_{i\oplus 1}^{-1}$ smoothly. The addition operation $\oplus$ on the index i is cyclic on the set $\{1, 2, 3\}$. By (38), the six boundary



(a)



(b)

FIGURE 5.4. Examples of boundary illegalities. (a) An example of a medial surface with an implied boundary that is not closed. (b) A swallowtail singularity on the implied boundary.

counterparts meet if

$$\mathbf{u}_i^1 = \mathbf{u}_{i\oplus 1}^{-1} \ , \tag{45}$$

since the r value is the same for all three subsurfaces at a seam point. Solving (39), (45) for $\mathbf{n}_i$ yields the following constraint on seam points:

$$\mathbf{n}_i = \frac{\nabla r_{i\oplus 2} - \nabla r_{i\oplus 1}}{\sqrt{1 - \|\nabla r_i\|^2}}. \tag{46}$$

At seam endpoints one of the angles between the three joining subsurfaces becomes π and the constraint (46) disappears. Applicable free edge constraints must still be satisfied.

The above constraints ensure that the implied boundary a is closed and connected surface but do not guarantee that it is non-singular. Fig. 5.4b shows an implied boundary forming a swallowtail singularity, which commonly occurs when the medial surface is left unconstrained. Singularities and invalid regions are detected using the Jacobian of the mapping in (38), given by

$$J^t = t\frac{\mathbf{b}_u^t \times \mathbf{b}_v^t \cdot \mathbf{u}^t}{\mathbf{m}_u \times \mathbf{m}_v \cdot \mathbf{n}} \ , \text{for } t = -1, 1. \tag{47}$$

At singularities $J^t = 0$ and there exists a region where $J^t < 0$. Therefore, singularities and points with reverse boundary orientation are eliminated by ensuring that $J^t > 0$ at each point on the medial surface.

When the generated boundary is convex or hyperbolic at a point, the inscribed sphere must have smaller radius than the radii of curvature corresponding to the negative principal curvatures. Otherwise, the inscribed sphere would cross the implied boundary. Hence the constraint

$$r < -\frac{1}{\kappa^t_{min}} \tag{48}$$

must be satisfied by a valid cm-rep.

In two dimensions the above set of constraints reduces to a set of simple requirements. At free endpoints $|dr/ds| = 1$, while $|dr/ds| < 1$ is enforced on the interior of medial curves. At shared endpoints where three medial curves meet at an angle to each other, the following constraint holds:

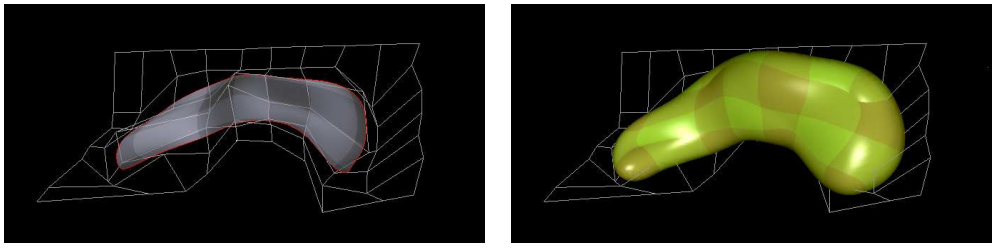
$$\frac{dr_i}{ds_i} = \mathbf{t}_{i\oplus 1} \cdot \mathbf{t}_{i\oplus 2} \ , \tag{49}$$

where the three tangent vectors are taken pointing into the shared endpoint. The Jacobian constraint from (47) simplifies to

$$\mathbf{b}^t_u \cdot \mathbf{t} > 0 \ , \tag{50}$$

and the radius of curvature constraint from (48) does not change.

5.3.3. Spline-Based Generative Model. Cubic b-splines are used to model the medial surface and the thickness scalar field because they provide local control and C^2 continuity. The formulation of a smooth medial manifold as a b-spline is similar in 2D and 3D. This subsection begins with the 3D case, which is presently limited to a single smooth medial surface.



(a) (b)

FIGURE 5.5. A manually constructed 3D b-spline medial of the hippocampus. (a) The medial surface and the control point polygon. (b) The implied boundary.

A smooth medial surface $(\mathbf{m}(u,v), r(u,v))$ is defined in terms of control points as follows:

$$\begin{aligned} \mathbf{m}(u,v) &= \sum_{i=0}^{d_1}\sum_{j=0}^{d_2} N_i^3(u)N_j^3(v)\bar{\mathbf{m}}_{ij} \ , \\ r(u,v) &= \sum_{i=0}^{d_1}\sum_{j=0}^{d_2} N_i^3(u)N_j^3(v)\bar{r}_{ij} \end{aligned} \tag{51}$$

where $(\bar{\mathbf{m}}_{ij}, \bar{r}_{ij}) \in \mathbb{R}^3 \times \mathbb{R}$ is a (d_1+1) by (d_2+1) array of control points that include both positional and thickness components. N_i are third order b-spline basis functions Piegl and Tiller [1996].

Medial b-splines must satisfy the legality constraints defined in section 5.3.2. The constraints ensuring that the generated boundary is closed are enforced differently in 2D and 3D.

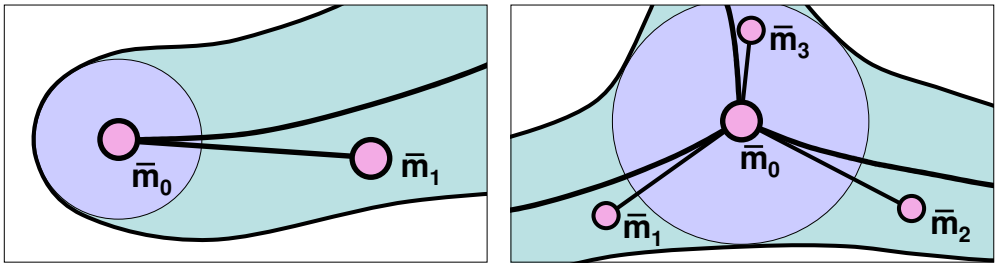
B-spline surfaces are defined on a quadrilateral mesh and thus have sharp corners. It is not practical (or even possible) to enforce the edge condition $\|\nabla r\| = 1$ on the edge of such a rectangular surface, as it would limit the range of objects that could be represented. Neither is is possible to define enforce the edge constraint along some predetermined curve on the b-spline surface, because the number of degrees of freedom afforded by the set of control points is finite and does not suffice for enforcing an arbitrary condition at each point of such curve.

However, the edge constraint can be satisfied if the edge itself is defined implicitly as the curve $\|\nabla r\| = 1$, and the portion of the b-spline surface where $\|\nabla r\| > 1$ is not considered to be part of the medial locus. The existence of the curve $\|\nabla r\| = 1$ on the b-spline surface is ensured by setting by setting $\bar{r}_{ij}$ to large negative values at the perimeter control points, and by keeping $\bar{r}_{ij}$ positive at the interior. The large difference in the values of $\bar{r}_{ij}$ near the perimeter causes causes the values of r_u and r_v to become large, in turn causing ∇r to exceed 1, as a consequence of (40). The small difference in the values of $\bar{r}_{ij}$ on the interior keeps $\|\nabla r\| < 1$ there. In order to describe objects of spherical topology, special care is taken to make sure that surface points where $\|\nabla r\| < 1$ form a single connected region, and hence, that the level curve is unique and not self-intersecting.

The level curve that defines the edge of a medial branch is computed using a numerical root finder in the process of sampling the spline. Whenever two consecutive samples have opposite signs of $\|\nabla r\| - 1$, Newton's method is applied along the vector connecting the two samples in parameter space. Fig. 5.3a is an example of a b-spline surface trimmed by the edge curve.

B-splines are an especially elegant representation for 2D cm-reps. Not only is it possible to represent medial structures with multiple connected branches, but also to incorporate all of the constraints on ∇r directly into the b-spline model. These constraints can be expressed as simple relationships between adjacent control points.

The knot sequences used to construct 2D b-splines have four repeating zero values at the beginning and four repeating unity values at the end. As a result, the b-spline behaves like a Bezier curve at the end points: it interpolates the terminal control points and the first derivatives of $(\mathbf{m}, r)$ at curve ends are linear combinations of the terminal and the next-to-terminal control points, and shown in Fig. 5.6a.



(a) (b)

FIGURE 5.6. Simple constraints on neighboring control points are used to ensure that the implied boundary is closed. (a) At free endpoints, the radius at the next-to-last control point is constrained. (b) At shared endpoints the control point $\bar{\mathbf{m}}_0$ is also shared and the radius value at the three neighboring control points is constrained.

At free endpoints, the constraint $|dr/ds| = 1$ is expressed in terms of the terminal control point $(\bar{\mathbf{m}}_0, \bar{r}_0)$ and the next control point $(\bar{\mathbf{m}}_1, \bar{r}_1)$:

$$\bar{r}_1 = \bar{r}_0 + \|\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_1\| \ . \tag{52}$$

At the points where three branches meet, the terminal control point of each branch is shared, as seen in Fig. 5.6b. The connectivity constraints from (49) are expressed in terms of the shared control point $(\bar{\mathbf{m}}_0, \bar{r}_0)$ and the next-to-last control points $(\bar{\mathbf{m}}_i, \bar{r}_i), i \in \{1, 2, 3\}$ of each branch:

$$\bar{r}_i = \bar{r}_0 + \|\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_i\| \frac{(\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_{i\oplus 1}) \cdot (\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_{i\oplus 2})}{\|\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_{i\oplus 1}\| \|\bar{\mathbf{m}}_0 - \bar{\mathbf{m}}_{i\oplus 2}\|} \ . \tag{53}$$

Hence, a closed and connected generated boundary is formed by expressing the next-to-terminal control values $\bar{r}_i$ in terms of near control points.

5.4. Parameter Estimation for Image Segmentation

Continuous medial models are used in image analysis applications following the deformable models framework. A template cm-rep is deformed to optimally fit the image data in the presence of a geometric prior term.

Template models are constructed manually by moving control points and adjusting their r value. In 2D, new branches are added by dividing an existing branch into two parts, and joining the two new branches with a third branch at a single control point. The user interface provides feedback when one of the constraints defined above is violated (although constraints on ∇r are enforced by the construction of the model, the non-singularity constraints must be checked after each modification of the model). The method often succeeds even if initialized with a template model that violates the non-singularity constraints. An example of a manually built 3D template model is shown in Fig. 5.5, and the accompanying movie `manual.mov` demonstrates model editing in 3D.

Template models are deformed to fit an image in an iterative process that minimizes an energy term in presence of constraints. The energy term consists of an image match component and a prior component. The image match component depends on the application. The convolution of the image with the Gaussian derivative kernel is used for gray-scale image segmentation. The aperture of the kernel is proportional to the local thickness of the object.

When fitting a model to a binary segmentation in 3D, the image match term is the mean square distance to the boundary of the binary object Danielsson [1980]. The volume overlap measure can be used instead, e.g. when fitting coarse models with few control points to highly detailed binary objects.

Regardless of the type of the image match function used, the image match energy component is integrated over the surface of the implied boundary. Integration is performed by sampling the medial spline at a sufficiently fine level of detail. Presently, the spline is sampled uniformly in parameter space. In 3D, the intersections of the trimming edge curve with the sampled grid are added to the set of regular samples collected on the interior of the medial surface.

The prior energy term favors models with low curvature. It imposes a penalty of the form

$$(|\bar{\mathbf{m}}_{i+1} - \bar{\mathbf{m}}_i| - |\bar{\mathbf{m}}_i - \bar{\mathbf{m}}_{i-1}|)^2 \tag{54}$$

on the neighboring control points. The penalty causes the control polygon to be relatively smooth. B-splines possess a minimum curvature property Farin [1995] which relates the smoothness of the control polygon to the smoothness of the spline.

The singularity constraints (47) are implemented as heavily weighted penalty functions and are integrated over the sampling grid. The constraints are computed numerically on the order of the sampling grid. Only the first derivatives of $(\mathbf{m}, r)$ are computed during the deformation process. Discrete sampling makes it possible for small local violations of the constraints to occur. The issues of sampling are addressed in the discussion section.

Energy minimization is performed using the $\mu + \lambda$ evolution strategy Michalewicz [1992] in two stages. In the similarity transform stage the template is scaled, rotated and translated to best match the image. During the deformation stage, small groups of adjacent control points are selected in random order and their values are optimized. Each group is optimized for several hundred iterations of the evolutionary algorithm, and several passes over the groups are made, until deformations cease to be significant.

The global non-deterministic optimization method was chosen because the image-driven energy function has many local optima. While there exists no mathematical bound on the number of iterations needed for the optimization to converge, the empirical bound in 3D is on the order of 1000 iterations per group of four control points. Convergence time is discussed in section 5.5. The algorithmic complexity is linear in the number of control points because a change to a group of control points only requires the reevaluation

of the energy function over a local portion of the object. The complexity is linear in the number of sample points as well.

5.5. Results

This section demonstrates the effectiveness of cm-reps in application to object representation and automatic image segmentation. In two dimensions, a model of a vertebra is automatically deformed to fit a slice of a CT image. In three dimensions, a model of the hippocampus is similarly fitted to a manually segmented MRI of the human brain.

5.5.1. 2D Vertebra Segmentation. The continuous m-rep method was used to automatically segment segment a lumbar vertebra in a 2D axial slice of the Visible Human abdominal CT subset The U.S. National Library of Medicine [2001], which is shown in Fig. 5.7a. Fig. 5.7b shows a manually constructed initial model of the vertebra with 10 connected branch curves. Notice that this model intentionally contains illegalities, shown as blue segments of the boundary. The result of automatic segmentation is shown in Fig. 5.7c. The optimization process corrected the illegalities and fitted the image boundaries closely. Fig. 5.7d shows a detail from fitted model, where in the top left corner the limitation the of the non-singularity constraint 50 on the Jacobian can be seen. The curvature constraint 48 was not enforced during this segmentation.

5.5.2. 3D Hippocampus Segmentation. A three-dimensional template model of the hippocampus was automatically fitted to binary segmentations of the left and right hippocampi in a magnetic resonance image, which is a part of a study on effects of schizophrenia on the shape of the hippocampus Gerig et al. [2001]. The template model of the right hippocampus was constructed manually by interactively moving control points in space. A binary image describing the average shape of the right hippocampus in a group of subjects was used for reference. This average shape was computed by Styner and Gerig [2001a] using the SPHARM description of all the right hippocampi in the schizophrenia

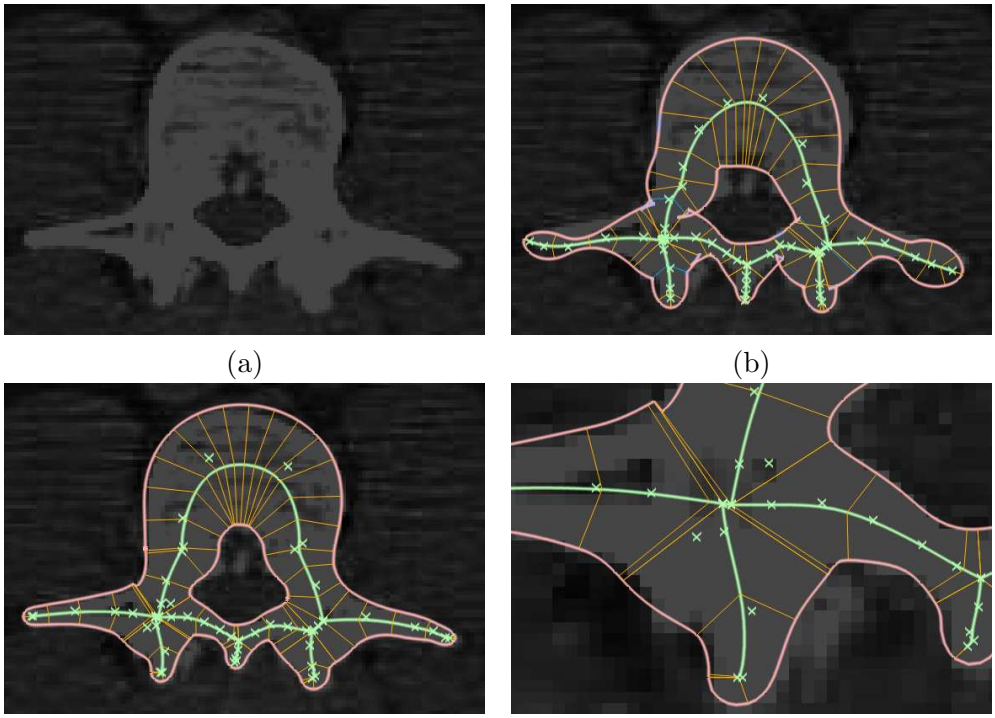
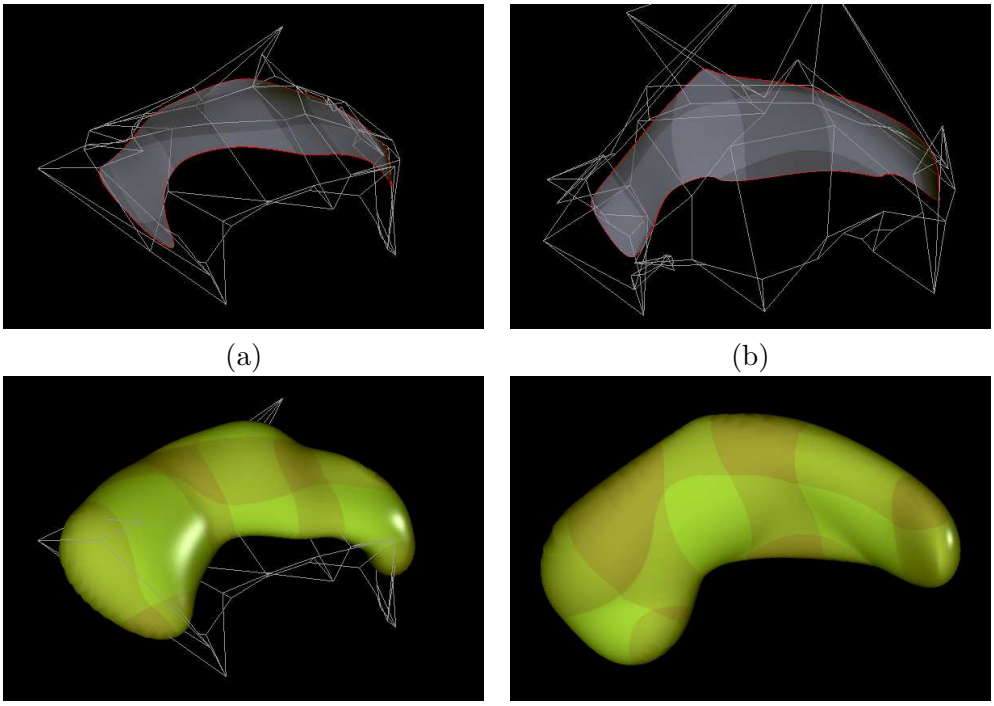


FIGURE 5.7. Automatic segmentation of a vertebra in an abdominal CT image. (a) A slice from the CT image. (b) A manually constructed model with illegalities. The yellow x-shaped marks denote the positions of control points. (c) A result of deforming the model to optimally fit the image. (d) A magnified detail of the model.

study. The template model of the left hippocampus was constructed by reflecting the right template across the midsagittal plane. The template model of the right hippocampus is shown in Fig. 5.5.

The similarity transform stage of the optimization was performed using a likelihood term that measures the volume overlap between the binary segmentation and the transformed template. A different likelihood term based on the mean square distance to the boundary was used for the deformation stage of the optimization. In order to compute mean square distance, the distance transform was computed on the binary image, and the result was blurred using a 1-voxel Gaussian kernel.



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

FIGURE 5.8. Automatically fitted models of the left and right hippocampi. (a) Left hippocampus: medial surface. (b) Right hippocampus: medial surface. (c) Left hippocampus: implied boundary. (d) Right hippocampus: implied boundary.

TABLE 5.2. Mean square distance from left and right hippocampal models to the binary segmentation.

	Left Model	Right Model
Template initialization	216.00	122.59
After similarity transform	4.34	4.55
After deformation	1.71	1.49

The optimization of a model with an order of 4000 sample points typically converges after 10-15 minutes on a 900 MHz Intel machine. The resulting models are shown in Fig. 5.8. Table 5.2 shows the mean square distances, in units of voxels, from the boundaries of the hippocampal models to the binary segmentations. Both the right and left hippocampus were fitted accurately. The accompanying movie `deform.mov` demonstrates 3D segmentation.

5.6. Discussion and Conclusions

5.6.1. Extensions to the Algorithm. Continuous m-reps have been used to represent 2D objects with branching medial topology and to represent 3D objects with a single medial surface. The extension of the method to cover 3D branching topology remains. The challenge lies in efficiently representing creases formed by seams in medial surfaces and in enforcing constraints along the creases implicitly. B-spline surfaces defined on a quadrilateral mesh can not represent seams efficiently because creases on these surfaces are formed by knot repetition. A crease extends across the entire surface and its image in the parameter space of the b-spline is parallel to the coordinate axes. Recent extensions to the b-spline paradigm Catmull and Clark [1978], Peters [1996] allow meshes of arbitrary topology to be used and promise to represent creases more efficiently. The 2D branching constraint (49) is expressed in terms of relationships between nearby control points. In 3D, the possibility of expressing the non-linear seam curve constraint (46) in terms of control points remains an open question.

In this dissertation, continuous m-reps are used to describe objects that have codimension 1 medial axes. Objects such as tubes with circular cross-sections in 3D and circles in 2D have medial axes of higher codimension. While objects with exact rotational symmetry are non-generic, some common biological structures, such as blood vessels in 3D, are almost perfectly tubular. Continuous m-reps fitted to such objects are sensitive to slight boundary perturbations and imaging noise. Curve-based medial representations, such as those used by **?**, provide a more stable description for tubular objects. A combined medial description could be defined; such a description could choose between surface-based and curve-based medial representations as dictated by image data.

The present implementation uses a uniform grid in b-spline parameter space to sample the cm-rep. Such sampling produces a non-uniform boundary grid, especially near edges of branches, where the medial-to-boundary mapping asymptotes. An adaptive

sampling scheme based on distances along the boundary would improve the robustness of segmentation and constraint enforcement.

Since the non-singularity constraints are checked only at the samples, small creases on the boundary are not always detected. An analytical expression of the constraints in terms of the b-spline control points needs to be derived in order to efficiently detect the boundary creases at all scales. Alternatively, a search for the maxima of the constraint function, although costly, would detect singularities independently of the sampling.

5.6.2. Comparison with Discrete M-Reps. One of the differences between continuous m-reps and discrete m-reps is the strictness with which the two representations conform to the rules of medial geometry. Continuous m-reps maintain a strict SAT relationship between the boundary and the medial locus. Discrete m-reps allow interpolation of both the boundary and the medial locus but only enforce the SAT relationship at the medial atoms. Differences in stability between the two methods remain to be analyzed.

The two methods also differ in the way they handle objects with branching medial topology. Continuous m-reps approximate the behavior of medial loci at branching points. Discrete m-reps do not attempt to simulate the way that medial branches seam and objects with multiple figures are represented by blending the boundaries of single figure m-reps. The blending makes is easy to model and deform complex 3D objects because a figure can be moved relative to another without recomputing the representation. Constructive solid geometry is possible with discrete m-reps.

Both representations provide an object-intrinsic metric on the space inside and around the object. This metric allows distance measures to take into account the local thickness of the object. For example, the aperture of the image intensity filter used in grayscale segmentation is proportional to the local thickness of the object. The intrinsic metric can be used to establish an object-centric coordinate system, which provides a framework

for fine-scale boundary perturbations within tolerances prescribed by the local thickness Joshi et al. [2002].

Recent advances in deformable segmentation, registration and shape analysis based on discrete m-reps can readily be extended to cm-reps. Applications that require arbitrary sampling of medial loci can benefit from the continuous representation. For instance, boundary sampling may be adjusted adaptively during segmentation, improving robustness and efficiency. Statistical shape analysis can benefit from cm-reps because correspondences between objects can be represented as a continuous mapping.

CHAPTER 6

Discussion and Conclusions

> He had been eight years upon a project for extracting sunbeams out of cucumbers, which were to be put into vials hermetically sealed, and let out to warm the air in raw, inclement summers.
>
> Jonathan Swift, *Gulliver's Travels*

6.1. Summary of Scientific Contributions

This section briefly summarizes the scientific contributions reported in Ch. 3, 4, and 5 of this dissertation. The section revisits each of the claims made in Sec. 1.3 and discusses where and how it has been addressed in the dissertation.

Claim 1. *New concept: A focus on locality in shape characterization*

In this dissertation, the term *locality* refers to the assumption that shape-altering processes, such as schizophrenia and other diseases, affect biological objects at one or few regions. For example, locality is exhibited by a small spherical tumor in an organ if over time the tumor displaces the organ's boundary in the region nearest to its center. Locality is a recurring theme in this dissertation, and it is the primary motivation behind the coarse-to-fine shape characterization framework described in Ch. 3 and the feature selection algorithm developed in Ch. 4.

In Ch. 3 locality is the basic reason for decomposing the shape characterization problem into a coarse-to-fine hierarchy of subproblems. This decomposition makes it possible

to find the regions where the shape-altering processes are most active, for example in context of a classification problem, where a comparison of the discriminating ability of the global coarse-scale subproblem and multiple local fine-scale subproblems is used to find such regions. Locality is also the basic assumption made in developing the shape-specific feature selection method, called window selection, in Ch. 4.

This dissertation does not set out to prove that shape features exhibit locality. Instead, it treats locality as reasonable assumption and develops methodology that takes advantage of it. However, the results of clinical data experiments reported in Ch. 4 seem to affirm that the assumption of locality is valid. The results reveal that the most significant differences between schizophrenic and healthy hippocampi occur in a few regions of the organ and suggest that locality is indeed exhibited by schizophrenia.

Claim 2. *New method: Using a multi-scale medial object representation for localizing and describing components of shape variability*

A new coarse-to-fine framework for two-dimensional shape analysis of objects was developed in Ch. 3. This approach decomposes the problems of shape characterization and shape-based classification into hierarchies of subproblems. Located at the top of such a hierarchy is a subproblem that examines global variability of objects at a coarse scale. At the bottom are subproblems that examine regions of objects locally at a fine scale. This hierarchy is conceptually similar to the process of human form perception, which examines objects at different levels of detail separately.

The coarse-to-fine framework uses discrete m-reps to describe objects in terms of their medial structure. It leverages the ability of m-reps to describe objects with an arbitrary level of tolerance, constructing a coarse-to-fine sequence of m-rep descriptions of each object. Statistical features used in the coarse-scale global shape characterization subproblem are derived from differences between elements of neighboring medial atoms in the coarse-level m-rep description. Features used in the fine-scale local subproblems

are derived from residual differences between medial atoms that describe the same region of the object at different scales. Residual features allow the global shape variability to be removed in the consideration of local variability.

An illustrative analogy can be made between coarse-to-fine shape features and wavelets. Wavelets are a kind of function decomposition basis whose components are localized both in time and in frequency. Coarse-to-fine residual features are similarly localized in terms of the range of places on an object that they describe and the range of levels of detail at which they describe these places. Figure 6.1a illustrates this analogy by showing the 'footprints' of different coarse-to-fine features in the space formed by location on one axis and level of detail on another. Coarse global features have a large footprint both in location and in level of detail, while the residual fine features have a small footprint of these dimensions. For comparison, Figs. 6.1b and 6.1c show the footprints of features that are not coarse-to-fine. For example, features derived from coefficients of a decomposition of objects into a basis of principal components of variability or into a basis of spherical harmonics have footprints that are bounded in level of detail but global in spatial location. On the other hand, features sampled directly from a dense geometric object representation, such as coordinate values of points sampled from the boundaries of objects, have footprints that are bounded in spatial location but global in terms of the level of detail.

The coarse-to-fine framework is designed to be used in the context of shape-based classification as a tool for localizing shape-altering processes responsible for differences between classes of biological objects. The ability to discriminate between classes is compared across the coarse-to-fine hierarchy of classification subproblems. If one or a few subproblems discriminate significantly better than others, one may conclude that the regions and levels of scale described by the features in those sub-problems are the ones most affected by the shape-altering process.

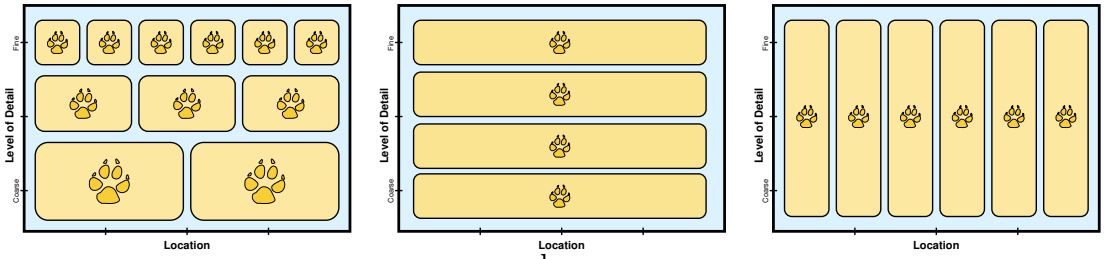


FIGURE 6.1. Feature 'footprints' in different shape characterization approaches. (a). In the approach described in Ch. 3 features have a hierarchical structure in the location-detail space: features that describe objects globally do so at a coarse scale, and spatially localized features describe objects at a fine scale. (b). The footprints of features computed as coefficients of global basis decompositions (e.g., eigenmodes, spherical harmonics) are bounded in level of detail but not in spatial location. (c). The footprints of features that measure local geometry, such as geometric properties sampled from a lattice are localized in space but not in level of detail.

The framework can also be used for estimating densities on m-rep parameters and for visualizing the principal components of variability in a sample of objects. The use of medial features makes it possible to decompose shape variability into geometrically qualitative components, such as variability in bending and variability in width and elongation, and to visualize these components separately.

The framework developed in Ch. 3 was tested using real and simulated shape data based on the shape of the corpus callosum. The experimental results verify the ability of the method to separately show components of shape variability that describe the bending of objects and the components that describe widening and elongation. In case of classification the results show that the method is able to detect the location of a fine-scale inter-class difference.

CLAIM 3. *New method: Selecting statistically relevant features in shape classification problems.*

Although the framework described in Ch. 3 is conceptually correct in using classification as a tool for localizing shape-altering processes, the approach that it takes for measuring the relative discrimination ability of classification subproblems is based on comparing p-values of mean difference tests and is not robust with respect to outliers. This approach does not provide an efficient way to compare relative discrimination ability of clusters of multiple local subproblems. A more theoretically sound approach to localizing shape-altering processes was presented in Ch. 4.

In Ch. 4, feature selection methodology from the machine learning literature was adapted to the problem of shape-based classification and localization of shape-altering processes. A novel algorithm, called window selection, was developed by extending an existing feature selection technique [Bradley and Mangasarian 1998], which uses support vector machine techniques and casts the problem of finding a subset of feature most relevant for discriminating between two classes as an energy minimization problem that can be solved efficiently using linear programming. The window selection algorithm incorporates locality, the assumption that shape-altering biological processes are localized to a region or a few regions of the object of interest, by including a heuristic that favors the selection of sets of features that represent clusters of neighboring locations in the geometric object representation. The window selection algorithm elegantly incorporates this locality heuristic into the linear programming formulation used in the feature selection method.

Claim 4. *Result: New feature selection method shown superior to a similar method in synthetic examples.*

Two kinds of experiments on synthetic data were performed in Ch. 4 in order to evaluate the performance of the window selection method as compared to the feature selection method on which it based and to classification without feature selection. The first kind of experiment is a multidimensional classification problem in which the features

are normally distributed, only a subset of the features is relevant for discrimination, and the relevant features are clustered together. When applied to this data, the window selection algorithm correctly identified the relevant set of features more frequently than the feature selection algorithm, and it produced classifiers that had better expected generalization performance than classifiers constructed using the feature selection algorithm or classifiers that use the entire set of features.

The second kind of experiment tested the performance of window selection in a simulated shape-based classification experiment. Two similar classes of objects that resemble the shape of the corpus callosum were generated, and a small circular outward displacement was applied to the objects in one of the classes. The center and radius of the displacement are distributed normally with small variance, causing the displacement to only affect a region of the corpus callosum shape. Statistical features were derived from a discrete representation of the corpus callosum boundaries. In this experiment two clusters of features are relevant for discrimination between classes. These clusters are located on opposite sections of the boundary nearest to the place where the circular displacement was embedded. When applied to these features, the window selection algorithm selected two clusters of features that overlap the clusters of relevant features, while the feature selection algorithm returned some irrelevant features located far away from the displacement location. The generalization ability of the classifier constructed in the space of features returned by window selection was marginally better than that of the classifier based on feature selection and significantly better when compared to a classifier constructed on the entire feature set.

Claim 5. *Application: New feature selection method is applied to clinical 3D hippocampus data.*

Chapter 4 described the application of the window selection and feature selection algorithms to clinical data from a study by Csernansky et al. [2002] of the effects of

schizophrenia on the shape of the hippocampus. The clinical data consists of dense meshes of the right and left hippocampi of 117 subjects, 52 of whom are schizophrenics, while the other 63 are matched healthy controls. The dense meshes representing the right and left hippocampi were partitioned into 80 patches each. A summary feature was computed for each patch by integration of the displacement of the patch from the mean hippocampal shape. These patch summary features were computed in order to make it possible for the feature sets returned by window and feature selection algorithms to cover a significant portion of the hippocampal surface, since the number of features selected by these algorithms is limited by the size of the training set.

The application of window and feature selection algorithms to the patch summary features in a leave-one-out cross-validation experiment resulted in similar correct classification rates. These rates did not differ significantly from rates achieved by classifiers that do not use feature selection. More interesting is an examination of the frequency with which the window and feature selection algorithms identified different patches on the hippocampal boundary in the leave-one-out experiments. Patches most frequently selected by the feature selection algorithm are scattered around the hippocampus, but clusters of patches selected by the window selection algorithm form a more structured pattern and clearly indicate that the head of the right hippocampus is a location where differences between schizophrenic and healthy hippocampi are most significant (see Fig. 4.7). This result confirms earlier conclusions made by Csernansky et al. [2002].

Claim 6. *New method: Representing 2D and 3D object geometry using a continuous medial representation based on splines.*

The ability to interpolate and resample medial atoms in discrete m-reps is necessary in order to produce residual features in the coarse-to-fine framework described in Ch. 3. A basic algorithm capable of such interpolation in two dimensions was presented in Sec. 3.2.1. The difficulty of extending this algorithm to three dimensions led to

the development of a new continuous medial representation that was described in Ch. 5. This continuous representation allows an object to be described at different levels of detail. These continuous descriptions can be arbitrarily sampled, and relationships between pairs of corresponding samples can be used as residual fine-scale features for shape characterization and classification.

Continuous m-reps are defined using a lattice of control points, which are tuples consisting of a position element and a radius element. The positions and radii are interpolated using cubic b-spline curves or surfaces, resulting in a continuous representation of a medial locus. A medial atom can be computed at each point on this continuous medial locus in terms of the zeroth and first order properties of locus at the point. The boundary nodes of the atoms computed over the interpolated medial locus form another locus that is continuous and closed under a certain set of constraints, which are discussed in Sec. 5.3.2. This locus describes the boundary of the object represented by a continuous m-rep. A nice mathematical property of the continuous m-reps is the fact that the medial locus defined by the m-rep is indeed the grassfire transform of the boundary locus generated by it.

In two dimensions, the spline-based representation is flexible enough to allow for branching medial loci that are needed to describe objects with multiple figures. In three dimensions the representation is presently limited to single-figure objects.

Claim 7. *Result: Continuous medial representation shown capable of describing and segmenting objects in 2D and 3D medical images.*

An automatic image segmentation algorithm used for discrete m-rep based segmentation was adapted to continuous m-reps in Ch. 5. The algorithm follows the deformable templates framework. It iteratively deforms a template continuous m-rep by updating the values of the b-spline control points, trying to optimize an objective function that

rewards match between edges in the image and the boundary generated by the continuous m-rep and penalizes control point configurations that violate constraints from Sec. 5.3.2 and cause the implied boundary to become invalid.

In two dimensions the algorithm was successfully applied to the segmentation of a vertebra in an axial CT slice of the human abdomen. The vertebra is a structure with a complex branching topology and the continuous m-rep template used to fit it contains multiple figures. In three dimensions, the algorithm was applied to a binary characteristic images of the hippocampus that had been segmented manually by experts from brain MRI images. In both applications of the segmentation algorithm, the resulting continuous m-reps accurately matched the boundaries of the objects in images.

6.2. Discussion and Future Work

> He told me, he did not doubt in eight years more, that he should be able to supply the governor's gardens with sunshine at a reasonable rate; but he complained that his stock was low, and intreated me to give him something as an encouragement to ingenuity, especially since this had been a very dear season for cucumbers.
>
> Jonathan Swift, *Gulliver's Travels*

Chapters 3,4, and 5, the three main chapters of this dissertation, are connected to each other conceptually rather loosely. In an analogy to software engineering, we can think of the coarse-to-fine framework developed in Ch. 3 as *Version 1.0* of a longer ongoing project. This early version is limited to two-dimensional single-figure objects. Its approach to localization of shape-altering processes is based on comparison of t-test p-values and does not truly identify object regions where the discrimination between classes of objects is most significant. The technique that it uses for medial atom interpolation does work for all possible configurations of medial atoms, etc.

Chapter 4 presented a comprehensive extension to the Version 1.0 in the areas of feature selection, which replaces the p-value based localization of shape-altering processes. The continuous medial representation discussed in Ch. 5 also is an extension to the Version 1.0 because it eliminates the need to interpolate medial atoms. In my view, the research accomplishments from Ch. 4 and 5 form components of *Version 2.0* of the shape characterization framework, my vision of which is presented and discussed in this section.

Version 2.0 will incorporate the coarse-to-fine ideas from Ch. 3, it will use feature selection, and it will rely on continuous m-reps for representing object geometry, for automatic image segmentation, and for deriving statistical features that are geometrically expressive. It will be able applicable to three-dimensional objects, even ones with branching medial topology. It will use an algorithm in the spirit of window selection presented in Ch. 5 for localizing regions of biological objects that are most significant for discrimination between classes of objects. However, in order for Version 2.0 to do all these wonderful things, its developers would have to solve a number of challenging problems.

In Version 2.0, three-dimensional biological objects, including objects with branching medial topology, will be represented by continuous m-reps. This capability requires an extension to the existing continuous m-rep methodology, which in its current state is not capable of representing curves, called shared edges, along which medial sheets in a medial locus with a branching topology meet. Such an extension has been attempted by P. Thomas Fletcher. His initial results were encouraging because he was able to model segments of shared edges. However, he has yet not been able to model the behavior of the medial locus at the special points located at the ends of the shared edge (i.e., the points of type A_1A_3 according to the classification by Giblin and Kimia [2000]). In order to model the shared edge, Fletcher had to switch to a higher order spline representation of the continuous radial scalar field, which allowed him to continuously constrain the

gradient of the field along the shared edge. It appears to be impossible to enforce these constraints using the cubic b-spline based definition of the continuous m-reps.

A change to a different spline representation may, with a considerable amount of effort, make it possible to model the shared edges and their endpoints. However, a higher order spline with with special conditions imposed on it is likely to lose the nice properties such as minimal curvature and local control, which make b-splines a popular choice for geometric modelling.

The technique used to represent multi-figure objects in discrete m-rep modeling can be extended to continuous m-reps. This technique is reminiscent of constructive solid geometry because it describes an object as a union of single-figure continuous m-reps whose boundaries are blended smoothly near their intersections, forming a closed continuous surface. The disadvantage of using this approach is that the Blum medial locus of the blend regions of the boundary would not coincide with the medial locus defined by the continuous m-rep as it does for boundaries of single-figure continuous m-reps.

Another approach to modeling three-dimensional multi-figure models would be to change the way in which the medial surfaces and their edges are treated conceptually. The current approach represents the medial surface explicitly as a b-spline patch, and it models the edges of the medial surface implicitly, as level curves of the gradient magnitude of the radius scalar field. The alternative is to explicitly represent the edges and intersections of edges of medial sheets as curves and points and to implicitly model the medial surfaces bounded by these edges. Conceptually this is analogous to drawing a sketch of the medial locus: an artist first draws the outline of an object and later fills in the interior. In fact, Giblin and Kimia [2000], who classify the different types of points, curves, and surfaces that generically occur on a three-dimensional medial locus, suggest a data structure for representing the relationships between these points, curves, and surfaces, and this data structure may be directly adaptable to this alternative type of continuous m-rep modeling. The difficulty of this approach lies in figuring out how

to control the interior portions of medial surfaces once their edges have been defined explicitly. In my opinion, this alternative approach may be the most promising, but it is also the one least explored.

In order to allow derivation of a coarse-to-fine hierarchy of statistical features, Version 2.0 will, like its predecessor, describe geometrical objects at multiple levels of detail. Each object will be represented by a sequence of continuous m-reps fitted to it with decreasing tolerance. The continuous m-rep methodology presented in Ch. 5 supports explicit specification of tolerance in roughly the same manner as it is specified for discrete m-reps in Version 1.0. However, since the inception of Version 1.0, the discrete m-rep methodology has evolved considerably, and better techniques for describing and modelling m-rep tolerance have been developed (for example, the technique of Lu et al. [2003b] based on Markov random fields, which is briefly summarized in Sec. 2.1.7). Version 2.0 should extend such techniques to continuous m-reps in order to improve the robustness and accuracy of segmentation.

In order to consistently describe populations of objects at various levels of detail using continuous m-reps, Version 2.0 will incorporate a methodology for ensuring good positional correspondence between instances of continuous m-reps that describe different objects and between instances that describe the same object with different tolerance. As summarized in Sec. 2.2.4, the shape characterization literature offers some solutions to the problem of correspondence for geometric representations that describe different objects. The following paragraphs consider the applicability of these solutions to continuous m-reps, in lieu of the special properties of the latter, e.g., their composition into figures and the alignment of their edges with the maxima of boundary curvature.

Let us begin by considering the problem of establishing correspondence between continuous m-reps that do not branch, i.e., consist of only one medial manifold. In this case, the problem of positional correspondence is continuous according to the categorization of Sec. 2.2.4: correspondence can be defined as a set of diffeomorphic mappings $\psi_1 \ldots \psi_n$

from a common domain to the parameter domains of n different instances of medial manifolds. Let us define optimal correspondence as a choice of $\psi_1 \ldots \psi_n$ that maximizes some measurement of match integrated over the common domain. For a point $\mathbf{u}$ in the common domain, we can compute the match in terms of the similarity between a set of medial manifold points associated with the parameter values $\psi_1(\mathbf{u}) \ldots \psi_n(\mathbf{u})$.

In some undocumented work with two-dimensional continuous medial representations of corpora callosa, I experimented with finding optimal correspondences between pairs of single-figure continuous m-reps by defining different measurements of similarity between medial points and by modelling the diffeomorphic transformations ψ_1, ψ_2 using cubic splines. In some experiments the similarity between two corresponding points on two different medial curves was computed as the weighted sum of squared differences between various geometrical properties derived at those points, including medial curvatures, values of the radius scalar field, and boundary curvatures at the boundary nodes associated with the medial points. In other experiments the overall match between two two medial curves reparameterized by diffeomorphic transformations ψ_1, ψ_2 was measured as the mutual information between point sets sampled from them densely. Once a measurement of match was selected, the optimal correspondence was found by deforming an initial set of spline control points defining ψ_1, ψ_2 in a way that minimized a Bayesian energy that consists of an likelihood term that integrates the match over the common domain and a prior term that penalizes large values of the derivative of ψ_1 and ψ_2. I found that my approach for establishing pairwise correspondence between medial curves consistently aligned local minima and maxima of one curve with the local minima and maxima of the other curve, in particular when applied to m-reps that describe the same object at difference levels of tolerance.

Based on my experience with two-dimensional single-figure medial correspondence, I envision two possible solutions to the three-dimensional problem of continuous correspondence between medial sheets. These two solutions are described in the following paragraphs.

The first potential solution comes from a realization that problem of establishing correspondence is analogous to non-rigid image registration. In both problems, one looks for a set of transformations that optimally align a set of measurements between a collection of objects. In the case of registration, the objects are images and the measurements are image intensities. In the case of correspondence, the objects are medial sheets and the measurements can be defined in many different ways and can be multi-valued. The corpus callosum callosum work described above is an example of applying registration techniques to the correspondence problem, and its extension to three dimensions is possible.

Many different measurements could be used as 'image intensities' for registration-like correspondence. For example, in the corpus callosum experiments, local measurements of medial curvature, radius, and curvature of the implied boundary were used. Such measurements could be taken at a scale, e.g., by convolving some original set of measurements over the parameter domain, thus extending the analogy to registration of image scale-spaces. The previously described notion of residues between object descriptions at different scales could be used in constructing these multiscale measurements: for example, the fine-scale difference between the true boundary of an object and the boundary implied by the medial representation of the object can be used to form a set of measurements, which would allow the 'pimples' and 'dimples' on the object's boundary to contribute to the computation of the optimal correspondence. The idea of using both boundary properties and medial properties leads us to understand that 'image intensities' need not be limited to measurements describing medial sheets and their implied boundaries. They could also describe the ambient space in which the object lies, for instance, as distances to anatomical landmarks defined in this space. Also, when working in the

context of greyscale image segmentation, one could define multi-valued measurements by sampling the intensities of the greyscale image along vectors that go from medial points to the associated points on the implied boundary. Because of the inherent difficulty of the problem of correspondence, a combination of all these types of measurements may be needed to produce good correspondence between medial sheets.

Another solution to establishing correspondence for continuous m-reps is to extend the minimum description length approach by Davies et al. [2002], which is summarized in Sec. 2.2.4. This approach casts the problem of correspondence in terms of minimizing the description length needed to transmit the geometrical information about a collection of similar objects within a given level of tolerance, e.g., the resolution of the image from which the objects are segmented. The difficulty in applying the MDL idea to continuous m-reps lies in its inherent dependency on discretization and the non-linear relationship between the regular sampling of object boundaries and regular sampling of medial loci. A regular discretization of a medial sheet within a particular level of tolerance does not necessarily result in a regular discretization of the implied boundary.

The search for optimal correspondence between single-figure continuous m-reps must take into account the fact that the ends of medial curves and edges of medial sheets are aligned with maxima of objects' boundary curvature and therefore inherently provide a a gross correspondence between these maxima. In the case of three-dimensional medial sheets, whose parameter domains are trimmed by a level curve that defines the edge of the medial sheet, special care must be taken to ensure that the mappings $\psi_1 \ldots \psi_n$ map the edges of the common parameter domain onto the trim curves.

In the case of multi-figure continuous m-reps the problem of correspondence becomes partially discrete and partially continuous. The discrete problem deals with making sure that corresponding medial sheets in different m-rep instances describe corresponding figures, e.g., when describing a population of hands, ensuring that the medial curve indexed

by number 2 always describes the index finger. The continuous problem deals with point-by-point similarity between corresponding points in the discretely corresponding medial sheets.

The statistical features in Version 2.0 will form a coarse-to-fine hierarchy. Like in Version 1.0, these features will describe relationships between neighboring locations on the continuous m-reps or between corresponding locations of m-reps that describe the same object at different levels of detail. Future research will show which local geometrical relationships will form good features. In some experiments that I have performed using 3D continuous m-reps, discrete approximations of curvature of the medial locus and the second derivative properties the of radial scalar field seemed to form a good set features. The sampling strategy used to compute finite sets of features from continuous medial sheets remains to be researched as well.

Version 2.0 will use the window selection methodology described in Ch. 5 as a tool for localization of shape-altering processes. The window algorithm will need to be adapted to handle features derived from the medial representation. Since these features use different types of measurements such as widening and bending to describe local geometry, the window selection algorithm needs to be able to handle what I call *co-located features*: different measurements that describe the same place in the geometric object representation. Co-located features were not dealt with in Ch. 5, where each location in an object representation was described by a single-valued feature, such as the distance from a boundary point to the origin in the synthetic shape examples in Sec. 4.3.2 and the patch summary feature computed as the average displacement from the mean shape in the hippocampus experiments in Sec. 4.4.3. One approach to handle having multiple features per location is to define the distance between any two co-located features to be equal to zero, causing windows to always include all of the co-located features together. However, this approach would cause irrelevant features to be returned by window selection, because the presence of a single relevant feature at a particular location will cause all of the features co-located

with it to be selected, regardless of their relevance. In a different approach, we can define the distance between co-located features to be infinite, causing these features never to be placed in the same window and resulting in the independent analysis of the features. The issue of co-located features requires further theoretical consideration.

Since the window selection methodology was developed and tested on features that describe objects at a single level of detail, an extension to handle coarse-to-fine feature hierarchies is necessary. The window selection could be adapted to hierarchical features if an appropriate distance function were defined on them. This distance function needs to reflect the fact that each location in an object is described by multiple coarse-to-fine features. The distance between features needs to combine a spatial component and a component that measures distance in scale, i.e., the number of levels of detail that separate two features.

In addition to defining a new distance function for coarse-to-fine features, Version 2.0 will most likely need to have the capability to apply window selection hierarchically. The need to do so arises from the limitation posed by the training sample size on the number of features that can be returned by the window selection algorithm. Recall that in Ch. 4 the set of features describing the boundary of the hippocampus had to be reduced to a much smaller set of patch summary features in order to allow window selection to identify as significant regions that cover a substantial portion of the boundary. If the hippocampi were to be described using a coarse-to-fine hierarchy of features, there would be even more features to select from, and the reduction to patch summary features would not be an option because it would lose the advantages of using a coarse-to-fine hierarchy.

In order to make window selection useful for large numbers of coarse-to-fine features, Version 2.0 needs to be able to perform window selection recursively in a hierarchical way. First, window selection would be applied to the few global coarse-scale features. Then the finer-scale features that describe the regions identified as relevant by the coarse-level window selection will be subjected to window selection, and the process will be

repeated until the finest features in the hierarchy are reached. This approach would reduce the number of features input into each application of window selection, eliminating the need to compute summary features before applying window selection. However, the effectiveness and theoretical validity of applying window selection hierarchically still need to be studied.

The list of open problems that lie on the patch to Version 2.0 is long and intimidating. Let us motivate the need for solving these problems by considering the potential benefits offered by the Version 2.0 framework. Version 2.0 would allow shape characterization and classification of complex anatomical objects, including subcortical structures including brain ventricles, abdominal structures, bones, and many other regions of the body that are affected by diseases. By performing shape analysis at a hierarchy of scales, Version 2.0 would be able to reduce its complexity is manner analogous to the way that the human visual system perceives the world: first at a coarse scale then locally at a fine scale. By representing object medially, Version 2.0 would be able to characterize shape variability in terms of geometrically descriptive components. And by combining descriptive features with feature selection, Version 2.0 might indeed be able to explain exactly where and how certain diseases affect human anatomy.

BIBLIOGRAPHY

Aha, D. W. and Bankert, R. L. A comparative evaluation of sequential feature selection algorithms. In *Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics*, pages 1–7. Springer-Verlag, New York, NY, 1995.

Arcelli, C. and Sanniti di Baja, G. Ridge points in euclidean distance maps. *Pattern Recognition Letters*, 13(4):237–243, 1992.

Attali, D., di Baja, G. S., and Thiel, E. Skeleton simplification through nonsignificant branch removal. *Image Processing and Communications*, 3(3–4):63–73, 1997.

August, J., Siddiqi, K., and Zucker, S. Ligature instabilities in the perceptual organization of shape. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1999.

Aylward, S. and Bullitt, E. Initialization, noise, singularities, and scale in height-ridge traversal for tubular object centerline extraction. *IEEE Transactions on Medical Imaging*, pages 61–75, 2002.

Aylward, S. *Continuous Mixture Modeling via Goodness-of-Fit Cores*. Ph.D. thesis, Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1997.

Bello, F. and Colchester, A. Measuring global and local spatial correspondence using information theory. In *In Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, volume 1496 of *Lecture Notes in Computer Science*, pages 964–973. Springer, 1998.

Biederman, I. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2):115–147, 1987.

Blanding, R., Brooking, C., Ganter, M., and Storti, D. A skeletal-based solid editor. In W. F. Bronsvoort and D. C. Anderson, editors, *Proceedings of the Fifth Symposium on Solid Modeling and Applications (SSMA-99)*, pages 141–150. ACM Press, New York, 1999. ISBN 1-58113-080-5.

Bloomenthal, J. and Shoemake, K. Convolution surfaces. *Computer Graphics (SIGGRAPH '91 Proceedings)*, 25(4):251–256, 1991.

Blum, H. and Nagel, R. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10(3):167–180, 1978.

Blum, H. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*. MIT Press, 1967.

Bookstein, F. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567–585, 1989.

Bradley, P. S., Mangasarian, O. L., and Street, W. N. Feature selection via mathematical programming. *INFORMS Journal on Computing*, 10:209–217, 1998.

Bradley, P. S. and Mangasarian, O. L. Feature selection via concave minimization and support vector machines. In *Proc. 15th International Conf. on Machine Learning*, pages 82–90. Morgan Kaufmann, San Francisco, CA, 1998.

Bradley, P., Mangasarian, O., and Rosen, J. Parsimonious least norm approximation. Technical Report 97-03, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, 1997.

Brady, M. and Asada, H. Smooth and local symmetries and their implementation. *International Journal of Robotics*, 3(3):36–61, 1984.

Braunstein, M. L., Hoffman, D. D., and Saidpour, A. Parts of visual objects: An experimental test of the minima rule. *Perception*, 18:817–826, 1989.

Brechbühler, C., Gerig, G., and Kübler, O. Parameterization of closed surfaces for 3-D shape description. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 61(154–170), 1995.

Burbeck, C. A. and Pizer, S. M. Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, 35(13):1917–1930, 1995.

Calabi, L. and Hartnett, W. Shape recognition, prairie fires, convex deficiencies and skeletons. *American Mathematical Monthly*, 75:335–342, 1968.

Calabi, L. On the shape of plane figures. Technical Report Technical Report PML No 60429 SR-1, Parke Mathematical Laboratories, 1965a.

Calabi, L. A study of the skeleton of plane figures. Technical Report Technical Report PML No 60429 SR-2, Parke Mathematical Laboratories, 1965b.

Carr, J. C., Mitchell, T. J., Beatson, R. K., Cherrie, J. B., Fright, W. R., McCallum, B. C., and Evans, T. R. Reconstruction and representation of 3D objects with radial basis functions. In S. Spencer, editor, *Proceedings of the Annual Computer Graphics Conference (SIGGRAPH-01)*, pages 67–76. ACM Press, New York, 2001.

Caselles, V., Kimmel, R., and Sapiro, G. Geodesic active contours. In *International Conference on Computer Vision*, pages 694–699. 1995.

Catmull, E. and Clark, J. Recursively generated b-spline surfaces on arbitrary topological meshes. *Computer-Aided Design*, 10:350–355, 1978.

Chen, D., Pizer, S. M., and Whitted, J. T. Using multiscale medial models to guide volume visualization. Technical Report TR99-014, University of North Carolina, Chapel Hill, 1999.

Christensen, G., Joshi, S., and Miller, M. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16:864–877, 1997.

Christensen, G., Miller, M., Vannier, M., and Grenander, U. Individualizing neuroanatomical atlases using a massively parallel computer. *Computer*, 29(1):32–38, 1996.

Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. Active shape models – their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

Cootes, T., Edwards, G., and Taylor, C. Active appearance models. In *European Conference on Computer Vision*, volume 2, pages 484–498. Freiburg, Germany, 1998.

Cootes, T., Taylor, C., Cooper, D., and Graham, J. Active shape models - their training and application. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 1(61):38–59, 1994.

Cover, T. and Campenhout, J. The possible orderings in the measurment selection problem. *IEEE Transactions Systems, Man and Cybernetics*, 7(9):657–661, 1977.

Csernansky, J. G., Wang, L., Jones, D., Rastogi-Cruz, D., Posener, J. A., Heydebrand, G., Miller, J. P., and Miller, M. I. Hippocampal deformities in schizophrenia characterized by high dimensional brain mapping. *Am. J. Psychiatry*, 159:2000–2006, 2002.

Csernansky, J., Joshi, S., Wang, L., Haller, J., Gado, M., Miller, J., Grenander, U., and Miller, M. Hippocampal morphometry in schizophrenia via high dimensional brain mapping. In *Proc. National Academy of Sciences*, volume 95, pages 11406–11411. 1998.

Culver, T. *Computing the medial axis of a polyhedron*. Ph.D. thesis, Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2000.

Damon, J. Ridges and cores for two dimensional images. *Journal of Mathematical Imaging and Vision*, 10:163–174, 1999.

Damon, J. Smoothness and geometry of boundaries associated to skeletal structures, 2002. Preprint, available on the Internet at http://www.math.unc.edu/Faculty/jndamon.

Danielsson, P. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

Davatzikos, C., Vaillant, M., Resnick, S., Prince, J., Letovsky, S., and Bryan, R. A computerized approach for morphological analysis of the corpus callosum. *Journal of Computer Assisted Tomography*, 20:207–222, 1995.

Davies, R., Twining, C., Cootes, T., Waterton, J. C., and Taylor, C. 3d statistical shape models using direct optimisation of description length. In *7th European Conference on Computer Vision (ECCV)*, volume 3, pages 3–21. 2002.

Delingette, H. Simplex meshes: a general representation for 3D shape reconstruction. In *Conf. on Computer Vision and Pattern Recognition (CVPR '94)*. 1994.

Dryden, I. and Mardia, K. *Statistical Shape Analysis*. John Wiley & Sons, New York, 1998.

Duda, R., Hart, P., and Stork, D. *Pattern Classification*. John Wiley & Sons, 2 edition, 2001.

Eberly, D., Gardner, R., Morse, B., Pizer, S., and Scharlach, C. Ridges for image analysis. *Journal of Mathematical Imaging and Vision*, 4:351–371, 1994.

Eberly, D. *Ridges in Image and Data Analysis*. Computational Imaging and Vision Series. Kluwer Academic Publishers, Dordrecht, NL, 1996.

Farin, G. *Curves and Surfaces for CAGD: a practical guide*. Academic Press, Inc, San Diego, CA, fourth edition edition, 1995.

Fletcher, P. T., Lu, C., Joshi, S., and Pizer, S. M. Gaussian distributions on lie groups and their application to statistical shape analysis, 2003a. In submission to International Conference on Information Processing in Medical Imaging.

Fletcher, P. T., Lu, C., and Joshi, S. Statistics of shape via principal component analysis on lie groups, 2003b. To appear in IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

Fridman, Y., Pizer, S. M., Aylward, S., and Bullitt, E. Segmenting 3d branching tubular structures using cores, 2003. In submission to Medical Image Computing and Computer-Assisted Intervention.

Fritsch, D., Chaney, E., Boxwala, A., McAuliffe, M., Raghavan, S., Thall, A., and J.R.D., E. Core-based portal image registration for automatic radiotherapy treatment verification. *International Journal of Radiation, Oncology, Biology, Physics (special issue on Conformal Therapy)*, 33(5):1287–1300, 1995a.

Fritsch, D., Eberly, D., Pizer, S., and McAuliffe, M. Stimulated cores and their applications in medical imaging. In *Information Processing in Medical Imaging 1995 (IPMI'95)*, pages 365–368. 1995b.

Fritsch, D., Pizer, S., Yu, L., Johnson, V., and Chaney, E. Segmentation of Medical Image Objects using Deformable Shape Loci. In *International Conference on Information Processing in Medical Imaging*, pages 127–140. Springer-Verlag, Berlin, Germany, 1997.

Furst, J. and Pizer, S. Marching optimal-parameter ridges: an algorithm to extract shapt loci in 3D images. In W. M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI '98*, volume 1496 of *Lecture Notes in Computer Science*, pages 780–787. Springer, New York, 1998.

Gascuel, M. An implicit formulation for precise contact modeling between flexible solids. In J. T. Kajiya, editor, *Computer Graphics (SIGGRAPH '93 Proceedings)*, volume 27, pages 313–320. 1993.

Gelston, S. M. and Dutta, D. Boundary surface recovery from skeleton curves and surfaces. *Computer Aided Geometric Design*, 12(1):27–51, 1995. ISSN 0167-8396.

Gerig, G., Styner, M., Shenton, M., and Lieberman, J. Shape versus size: Improved understanding of the morphology of brain structures. In W. Niessen and M. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2208, pages 24–32. Springer, New York, 2001.

Giblin, P. J. and Brassett, S. A. Local symmetry of plane curves. *Amer. Math. Monthly*, 92:689–707, 1985.

Giblin, P. and Kimia, B. A formal classification of 3D medial axis points and their local geometry. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 566–573. 2000.

Golland, P., Fischl, B., Spiridon, M., Kanwisher, N., Buckner, R. L., Shenton, M. E., Kikinis, R., Dale, A. M., and Grimson, W. E. L. Discriminative analysis for image-based studies. In *Medical Image Computing and Computer-Assisted Intervention*

(MICCAI), volume 1, pages 508–515. Springer, 2002.

Golland, P., Grimson, W., and Kikinis, R. Statistical shape analysis using fixed topology skeletons: Corpus callosum study. In *International Conference on Information Processing in Medical Imaging*, LNCS 1613, pages 382–388. Springer Verlag, 1999.

Golland, P., Grimson, W., Shenton, M., and Kikinis, R. Deformation analysis for shaped based classification. In *International Conference on Information Processing in Medical Imaging*. Springer-Verlag, Berlin, Germany, 2001.

Gower, J. Generalized procrustes analysis. *Psychometrika*, 40:33–51, 1975.

Grenander, U. *Pattern Synthesis: Lectures in Pattern Theory, vol. I*. Springer-Verlag, New York, 1976.

Grenander, U. *Pattern Synthesis: Lectures in Pattern Theory, vol. II*. Springer-Verlag, New York, 1978.

Grenander, U. *Regular Structures: Lectures in Pattern Theory, vol. III*. Springer-Verlag, New York, 1981.

Haller, J., Banerjee, A., Christensen, G., Gado, M., Joshi, S., Miller, M., Sheline, Y., Vannier, M., and Csernansky, J. Three-dimensional hippocampal MR morphometry by high-dimensional transformation of a neuroanatomic atlas. *Radiology*, 202:504–510, 1997.

Highley, J. R., Esiri, M. M., McDonald, B., Cortina-Borja, M., Herron, B., and T.J., C. The size and fibre composition of the corpus callosum with respect to gender and schizophrenia: a post-mortem study. *Brain*, 122:99–110, 1999.

Hoffmann, C. M. and Vermeer, P. J. Validity determination for MAT surface representation. In G. Mullineux, editor, *Proceedings of the 6th IMA Conference on the Mathematics of Surfaces (IMA-94)*, volume VI of *Mathematics of Surfaces*, pages 249–266. Clarendon Press, Oxford, 1996.

Hoffman, D. D. and Richards, W. A. Parts of recognition. *Cognition*, 18:65–96, 1984.

Hohne, K. H., Bomans, M., Riemer, M., Schubert, R., Tiede, U., and Lierse, W. A 3D anatomical atlas based on a volume model. *IEEE Computer Graphics Applications*, 12(4):72–78, 1992.

Hubel, D. H. and Wiesel, T. N. Functional architecture of macaque monkey visual cortex. In *Proceedings of the Royal Society of London*, volume 198 of *B*, pages 1–59. 1977.

Igarashi, T., Matsuoka, S., and Tanaka, H. Teddy: A sketching interface for 3d freeform design. *Proceedings of SIGGRAPH 99*, pages 409–416, 1999. ISBN 0-20148-560-5. Held in Los Angeles, California.

Jain, A. K., Duin, R. P. W., and Mao, J. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, 2000.

Jain, A. and Zongker, D. Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):153–158, 1997.

Jebara, T. S. and Jaakkola, T. S. Feature selection and dualities in maximum entropy discrimination. In *Uncertainty in Artificial Intelligence: Proceedings of the Sixteenth Conference (UAI-2000)*, pages 291–300. Morgan Kaufmann Publishers, San Francisco, CA, 2000.

John, G. H., Kohavi, R., and Pfleger, K. Irrelevant features and the subset selection problem. In *International Conference on Machine Learning*, pages 121–129. 1994.

Jonker, P. and Vossepoel, A. On skeletonization algorithms for 2, 3 .. n dimensional images. In D. Dori and A. Bruckstein, editors, *Shape, Structure and Pattern Recognition*, pages 71–80. World Scientific, Singapore, 1995.

Joshi, S., Grenander, U., and Miller, M. On the geometry and shape of brain submanifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1317–1343, 1997.

Joshi, S., Pizer, S., Fletcher, P. T., Thall, A., and Tracton, G. Multi-scale 3-d deformable model segmentation based on medical description. In *International Conference on Information Processing in Medical Imaging*. Springer-Verlag, Berlin, Germany, 2001.

Joshi, S., Pizer, S., Fletcher, P., Yushkevich, P., Thall, A., and Marron, J. Multi-scale deformable model segmentation and statistical shape analysis using medial descriptions. *Invited submission to IEEE-TMI*, page t.b.d., 2002.

Joshi, S. *Large Deformation Diffeomorphisms and Gaussian Random Fields for Statistical Characterization of Brain Sub-Manifolds*. Ph.D. thesis, Washington Univ., St. Louis, MO, 1997.

Karypis, G. and Kumar, V. *MeTiS – A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices – Version 4.0*. University of Minnesota, 1998.

Katz, R. *Form Metrics for Interactive Rendering via Figural Models of Perception*. Ph.D. thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2002.

Kelemen, A., Székely, G., and Gerig, G. Elastic model-based segmentation of 3D neuroradiological data sets. *IEEE Transactions on Medical Imaging*, 18:828–839, 1999.

Keller, R. *Generic Transitions of Relative Critical Sets in Parametrized Families with Applications to Image Analysis*. Ph.D. thesis, Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1999.

Kendall, D. G. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of London Mathematical Society*, 16:81–121, 1984.

Kimia, B., Tannenbaum, A., and Zucker, S. Shape, shocks, and deformations i: The components of two-dimensional shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.

Kira, K. and Rendell, L. The feature selection problem: Traditional methods and a new algorithm. In *Tenth National Conference Conference on Artificial Intelligence (AAAI-92)*, pages 129–134. MIT Press, 1992.

Kohavi, R. and John, G. H. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324, 1997.

Kotcheff, A. and Taylor, C. Automatic construction of eigenshape models by genetic algorithm. In *International Conference on Information Processing in Medical Imaging*, pages 1–14. 1997.

Lee, D. Medial axis transformation of a planar shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(4):363–369, 1982.

Leyton, M. Symmetry-curvature duality. *Computer Vision, Graphics, and Image Processing*, 37(3):327–341, 1987.

Leyton, M. *Symmetry, causality, mind*. The MIT Press, Cambridge MA, 1992.

Lorensen, W. E. and Cline, H. E. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics (SIGGRAPH '87 Proceedings)*, 21:163–169, 1987.

Lu, C., Pizer, S. M., and Joshi, S. A markov random field approach to multi-scale shape analysis, 2003a. To appear in Scale Space 2003.

Lu, C., Pizer, S., and Joshi, S. Multi-scale shape modeling by markov random fields, 2003b. In submission to International Conference on Information Processing in Medical Imaging.

Malandain, G., Bertrand, G., and Ayache, N. Topological segmentation of discrete surfaces. *International Journal of Computer Vision*, 10(2):183–197, 1993.

Markosian, L., Cohen, J. M., Crulli, T., and Hughes, J. Skin: a constructive approach to modeling free-form shapes. *Computer Graphics*, 33(Annual Conference Series):393–400, 1999. ISSN 0097-8930.

Matheron, G. Examples of topological properties of skeletons. In J. Serra, editor, *Image Analysis and Mathematical Morphology Volume 2: Theoretical Advances*, pages 217–238. Academic Press, London, 1998.

McInerney, T. and Terzopoulos, D. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

Michalewicz, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, Berlin, 1992. ISBN 3-540-60676-9.

Miller, J. *Relative critical sets in $\mathbb{R}^n$ and applications to image analysis*. Ph.D. thesis, Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1998.

Morse, B., Pizer, S., Puff, D., and Gu, C. Zoom-invariant vision of figural shape: Effects on cores of image disturbances. *Computer Vision and Image Understanding*, 69:72–86, 1998.

Mumford, D. Pattern theory: A unifying perspective. In D. C. Knill and W. Richards, editors, *Perception as Bayesian Inference*, pages 25–62. Cambridge Universiry Press, 1996.

Nackman, L. R. *Three-Dimensional Shape Description Using the Symetric Axis Transform*. Ph. d. thesis, University of North Carolina at Chapel Hill, 1982. Department of Computer Science.

Näf, M., Kübler, O., Kikinis, R., Shenton, M., and Székely, G. Characterization and recognition of 3D organ shape in medical image analysis using skeletonization. In *Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 139–150. IEEE Computer Society, 1996.

Näf, M. *Voronoi Skeletons: a semicontinous implementation of the 'Symmetric Axis Transform' in 3D space*. Ph.D. thesis, ETH Zürich , Commmunication Technology

Institue, Image Analysis Group IKT/BIWI, 1996.

Narendra, P. M. and Fukunaga, K. A branch and bound algorithm for feature subset selection. *IEEE Transactions on Computers*, 26(9):917–922, 1977.

Netter, F. H. *Atlas of Human Anatomy*. Icon Learning Systems, second edition, 1997.

Ogniewicz, R. L. and Kübler, O. Hierarchic Voronoi skeletons. *Pattern Recognition*, 28(3):343–359, 1995.

Ogniewicz, R. *Discrete Voronoi Skeletons*. Hartung-Gorre Verlag, 1993.

Orban, G., Kato, H., and Bishop, P. End-zone region in receptive fields of hypercomplex and other striate neurons in the cat. *Journal of Neurophysiology*, 42:818–832, 1979.

Peters, J. Curvature continuous spline surfaces over irregular meshes. *Computer Aided Geometric Design*, 13(2):101–131, 1996.

Piegl, L. and Tiller, W. *The NURBS Book*. Springer-Verlag, New York, New York., 2 edition, 1996.

Pizer, S. M., Fletcher, P. T., Thall, A., Styner, M., Gerig, G., and Joshi, S. Object models in multiscale intrinsic coordinates via m-reps. *image and vision computing*, 21(1):1–16, 2003.

Pizer, S. M., Siddiqi, K., Székely, G., Damon, J. N., and Zucker, S. W. Multiscale medial loci and their properties, 2002. In submission to International Journal of Computer Vision.

Pizer, S., Eberly, D., Morse, B., and Fritsch, D. Zoom-invariant figural shape: The mathematics of cores. *Computer Vision and Image Understanding (CVIU '98)*, 69:55–71, 1998.

Pizer, S., Fletcher, T., Fridman, Y., Fritsch, D., Gash, A., Glotzer, J., Joshi, S., Thall, A., Tracton, G., Yushkevich, P., and Chaney, E. *Deformable M-Reps for 3D Medical Image Segmentation*. In Review, ftp://ftp.cs.unc.edu/pub/users/nicole/defmrep3d.final.pdf, 2000.

Pizer, S., Fritsch, D., Yushkevich, P., Johnson, V., and Chaney, E. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18:851–865, 1999.

Presse, W., Flannery, B., Teukolsky, S., and Wetterling, W. *Numerical Recipies in C*. Press Syndicate of the University of Cambridge, 1988.

Rangarajan, A., Chui, H., and Bookstein, F. The SoftAssign procrustes matching algorithm. In *International Conference on Information Processing in Medical Imaging*, volume 1230 of *LNCS*, pages 29–42. 1997.

Reyment, R., Blackith, R., and Campbell, N. *Multivariate Morphometrics*. Academic Press, New York, 2nd edition, 1984.

Rock, I. and Linnett, C. Is a perceived shape based on its retinal image? *Perception*, 22(1):61–76, 1993.

Schmitt, M. Some examples of algorithms analysis in computational geometry by means of mathematical morphology techniques. In J. D. Boissonnat and J. P. Laumond, editors, *Geometry and Robotics*, volume 291 of *Lecture Notes in Computer Science*, pages 225–246. Springer-Verlag, Berlin, 1989.

Serra, J. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.

Sethian, J. *Level Set Methods: Evolving Interfaces in Gometry, Fluid Mechanics, Computer Vision, and Material Sciences*. Cambridge University Press, 1996.

Shenton, M., Gerig, G., McCarley, R., Szekely, G., and Kikinis, R. Amygdala-hippocampus shape differences in schizophrenia: The application of 3D shape models to volumetric mr data. *Psychiatry Research Neuroimaging*, pages 15–35, 2002.

Sherbrooke, E., Patrikalakis, N., and E.Brisson. An algorithm for the medial axis transform of 3d polyhedral solids. *IEEE Transactions on Visualisation and Computer Graphics*, 2(1):44–61, 1996.

Sherstyuk, A. Shape design using convolution surfaces. In *Proceedings of Shape Modeling International '99*. 1999.

Siddiqi, K., Ahokoufandeh, A., Dickinson, S., and Zucker, S. Shock graphs and shape matching. In *International Conference on Computer Vision*. 1998.

Siddiqi, K., Ahokoufandeh, A., Dickinson, S., and Zucker, S. Shock graphs and shape matching. *International Journal of Computer Vision*, 1(35):13–32, 1999a.

Siddiqi, K., Bouix, S., Tannenbaum, A., and Zucker, S. W. The hamilton-jacobi skeleton. In *Proc. Computer Vision*, volume 2, pages 828–834. IEEE, 1999b.

Staib, L. and Duncan, J. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, 1992.

Stetten, G. D. and Pizer, S. M. Automated identification and measurement of objects via populations of medial primitives, with application to real time 3D echocardiography. *Lecture Notes in Computer Science*, 1613:84–97, 1999.

Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, 36:111–147, 1974.

Storti, D. W., Turkiyyah, G. M., Ganter, M. A., Lim, C. T., and Stat, D. M. Skeleton-based modeling operations on solids. In *SMA '97: Proceedings of the Fourth Symposium on Solid Modeling and Applications*, pages 141–154. ACM, 1997. Held May 14-16, 1997 in Atlanta, Georgia.

Styner, M., Gerig, G., Lieberman, J., Jones, D., and Weinberger, D. Statistical shape analysis of neuroanatomical structures based on medial models. *Medical Image Analysis (MEDIA, Elsevier)*, 2003. In print.

Styner, M. and Gerig, G. Medial models incorporating object variability for 3D shape analysis. In *International Conference on Information Processing in Medical Imaging*, pages 503–516. Springer-Verlag, Berlin, Germany, 2001a.

Styner, M. and Gerig, G. Three-dimensional medial shape representation incorporating object variability. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 651–656. 2001b.

Styner, M. *Combined Boundary-Medial Shape Description of Variable Biological Objects*. Ph.D. thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2001.

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. *Medical Image Analysis*, 1(1):19–34, 1996.

Székely, G. *Shape Characterization by Local Symmetries*. Habilitationsschrift, Institut fur Kommunikationstechnik, Fachgruppe Bildwissenschaft, ETH Zürich, 1996.

Tagare, H., O'Shea, D., and Rangarajan, A. A geometric criterion for shape-based non-rigid correspondence. In *International Conference on Computer Vision*, pages 434–439. 1995.

Tagare, H. Non-rigid curve correspondence for estimating heart motion. In *International Conference on Information Processing in Medical Imaging*, pages 489–494. 1997.

Talairach, J. and Tournoux, P. *Co-planar stereotaxic atlas of the human brain*. Verlag, Stuttgart, 1988.

Teixeira, R. C. *Curvature Motions, Medial Axes and Distance Transforms*. Ph.D. thesis, Harvard University, Cambridge, Massachusetts, 1998.

Thall, A., Fletcher, P. T., and Pizer, S. M. Deformable solid modeling using sampled medial surfaces: a multiscale approach. Technical report TR00-005, University of North Carolina, Chapel Hill, 2000.

Thall, A. Fast C^2 interpolating subdivision surfaces using iterative inversion of stationary subdivision rules. Technical report TR02-001, University of North Carolina, Chapel Hill, 2002.

The U.S. National Library of Medicine. The visible human project. 2001.

Thompson, P. and Toga, A. A framework for computational anatomy. *Computing and Visualization in Science*, 5:13–34, 2002.

Tsao, Y. and Fu, K. Stochastic skeleton modeling of objects. *Computer Vision, Graphics, and Image Processing, 1984*, 25:348–70, 1984.

Van Essen, D., Drury, H., Joshi, S., and Miller, M. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 95:788–795, 1998.

Van Essen, D. and Drury, H. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Journal of Neuroscience*, 17:7079–7102, 1997.

Van Essen, D. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*, 385:313–318, 1997.

Vapnik, V. N. *The Nature of Statistical Learning Theory*. Springer Verlag, 1995.

Vapnik, V. N. *Statistical Learning Theory*. John Wiley & Sons, 1998.

Vermeer, P. J. *Medial Axis Transform to Boundary Representation Conversion*. Ph. d. thesis, Purdue University, 1994.

Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., and Vapnik, V. Feature selection for SVMs. In *Advances in Neural Information Processing Systems 13*, pages 668–674. MIT Press, 2001.

Wunderling, R. *Paralleler und Objektorientierter Simplex-Algorithmus*. Ph.D. thesis, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, 1996. ZIB technical report TR 96-09.

Yuille, A. and Leyton, M. 3d symmetry-curvature duality theorems. *CVGIP*, 52:124–140, 1990.

Yushkevich, P., Fletcher, P. T., Joshi, S., Thall, A., and Pizer, S. M. Continuous medial representations for geometric object modeling in 2D and 3D. Technical report TR02-003, University of North Carolina, Chapel Hill, 2002.

Yushkevich, P., Fletcher, P. T., Joshi, S., Thall, A., and Pizer, S. M. Continuous medial representations for geometric object modeling in 2D and 3D. *image and vision computing*, 21(1):17–28, 2003.

Yushkevich, P., Pizer, S. M., Joshi, S., and Marron, J. S. Intuitive, localized analysis of shape variability. In *International Conference on Information Processing in Medical Imaging*, pages 402–408. Springer-Verlag, Berlin, Germany, 2001.

Zhu, S. and Yuille, A. FORMS: A flexible object recognition and modeling system. *International Journal of Computer Vision*, 20(3):187–212, 1996.