SHAPE MODELING FRAMEWORK FOR BRAIN AND FACIAL IMAGE
ANALYSIS

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

QING HE

Dr. Ye Duan, Dissertation Supervisor

MAY 2011

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

FAMILY TYPE AND INCIDENCE
OF CHILDHOOD DEPRESSION

presented by Qing He,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Professor Ye Duan

_____

Professor Dong Xu

_____

Professor Jianlin Cheng

_____

Professor James Keller

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

x

LIST OF TABLES

SHAPE MODELING FRAMEWORK FOR BRAIN AND FACIAL IMAGE ANALYSIS

Qing He

Dr. Ye Duan, Dissertation Supervisor

ABSTRACT

The field of medical image analysis has greatly influenced the practice of neuroscience. Many studies aim to find new disease related anatomical characteristics based on the analysis of Magnetic Resonance images (MRI). Image segmentation is usually a fundamental step, in which brain structures of interest are delineated from MRI. Neuroscientists then categorize segmentations of anatomical structures by attributes such as shape, size, or location. Besides medical images, biometric images such as facial images are gaining more and more attention from clinical research because certain biometric features are also related to neurological diseases. Similar to brain morphology analysis, facial features are collected for studies testing the hypothesis that the disease alters the facial shape.

Despite the rapid development of image analysis technologies, most clinical practice still relies on laborious manual. The aim of this thesis is to develop a shape modeling framework for brain and facial image analysis, for the purpose of clinical practice. Specifically, it focuses on the following three problems: (1) brain structure segmentation from MRI, where several deformable model based segmentation methods are proposed; (2) quantitative analysis of brain morphology, where methods for directly comparing 3D shapes of brain structures are proposed, and clinical applications are demonstrated; (3) facial morphology analysis, where an automatic facial feature localization method is proposed, and applied to the shape analysis of the eyes.

1. Introduction

1.1 Overview

The field of medical image analysis has greatly influenced the practice of neuroscience. Many studies aim to find new disease related anatomical characteristics in order to increase the reliability of diagnosing the illness or improving the effectiveness of treatment methods against the disease. For example, many studies of autism suggest that the size of the corpus callosum is altered by the disease [Piven et al., 1997; Hardan et al., 2000; Hardan et al., 2009; Keary et al. 2009; Frazier and Hardan 2009]. This might explain why some autistic patients have impaired communication and cognitive ability. These studies are often based on the analysis of Magnetic Resonance images (MRI). Image segmentation is usually a fundamental step, in which brain structures of interest are delineated from MRI for the subsequent analysis. Neuroscientists then categorize segmentations of anatomical structures by attributes such as shape, size, or location. These subject specific characterizations are collected for studies testing the hypothesis that the disease alters an anatomical structure.

Besides medical images, biometric images such as facial images are gaining more and more attention from clinical research because certain biometric features are also related to neurological diseases. For example, the presence of a specific facial pattern in patients with a genetic syndrome indicates phenotypic expression of this disease. Similar to brain morphology analysis, characterizations of facial surfaces are collected for studies testing the hypothesis that the disease alters the facial shape.

Despite the rapid development of image analysis technologies, most clinical practice still relies on manual operations which are labor intensive and subject to human factors. There is a need to bridge the gap between computational technology and clinical applications. The aim of this thesis is to develop a shape modeling framework for brain and facial image analysis, for the purpose of clinical practice. Specifically, it focuses on the following three problems: brain structure segmentation from MRI, quantitative analysis of brain morphology and facial morphology analysis.

Our major clinical application is the autism spectrum disorder (ASD) which is of great concern to the practicing pediatrician since the reported incidence has increased markedly over the past decade [Schaefer et al., 2008; Muhle et al., 2004]. Autism is a heterogeneous group of disorders united by common

behavioral phenotype, such as impairments in social interaction, impairments in verbal and non-verbal communication, restricted repetitive and stereotyped patterns of behavior, interests and activities [American Psychiatric Association, 2000; Muhle et al., 2004; Klauck et al., 2006]. It has so far defied attempts to discover either common genes or unifying brain morphology. It is well accepted that several susceptibility genes are interacting together with a complex mode of inheritance leading to the typical phenotypes of ASD, yet the genetic origins of ASD remain largely unknown because of the complex behavioral phenotypes and multigenic etiology of this disorder [Klauck et al., 2006; Lintas and Persico 2009; Losh et al., 2008].

Identification of specific and consistent regions of brain and facial abnormalities in autism will increase the basic knowledge of phenotypes in autism and how they differ from normal. Genetic heterogeneity in autism often prevents the finding of consistent phenotypes. A new definition of autism subtypes has been proposed [Miles et al., 2005] which divides autism into essential autism and complex autism. Complex autism is defined by the presence of a significant number of physical anomalies and/or macrocephaly. The remainders are defined as having essential autism. The essential and complex subgroups are relatively easy to identify clinically, and appear to differ in their outcomes, recurrence risks, sex ratios, and family histories. Most of the brain morphology analysis in this thesis focuses on essential autism only and exclude complex autism. This is partly due to the limitation of time and data availability. Moreover, essential autism comprises 80% of the ASD population [Miles et al., 2005] and is more homogeneous than the entire ASD population, so limiting the study to essential autism will reduce the heterogeneity of the subjects while revealing the characteristics of the majority of ASD.

We have also observed clinically that children in a core population of autism have a common facial gestalt, which can be described as physically pleasing with wide spaced, deep set eyes and a broad forehead. This group is named as Midline Facial Phenotype (MFP) group. Study of this facial gestalt is considered potentially informative in identifying the core population of autism which is a more homogeneous subgroup compared to the rest of the population. The phenotype is subtle and not easily categorized by standard measurements, thus a computer based procedure which is not subjective to human operators is in demand. This is the major goal of the facial morphology analysis in this thesis.

1.2. Brain MRI Segmentation

In recent years, medical image segmentation has become a standard technique for visualizing structures of the human brain as well as performing various types of volumetric and shape comparisons among these structures. Since the introduction of medical image segmentation, many methods have been implemented for brain structure segmentation MRI. These methods can be categorized into manual, semi-automatic [Yusgkevich et al., 2006; McInerney et al., 2006; Stalling and Hege, 1996; Liang et al., 2006], and fully automatic [Lee et al., 2000; McInerney et al., 2002; Hamarneh and McIntosh, 2005; Hata et al., 2000; Bello et al., 2007] methods. Manual segmentation is tedious, requires much attention to detail, and the results are not reproducible. On the other hand, fully automatic methods require no user interaction and are completely reproducible for the same data, but the reliability of these methods is often an issue because it solely depends on the method and no human knowledge can be of any help. Therefore, semi-automatic methods have become preferred for certain types of medical image segmentation, because doctors can control over the segmentation through user interaction. This thesis aims at developing effective segmentation methods that minimize the laborious user operation while keeping sufficient flexibility for clinical use.

Brain MRI segmentation can be classified into tissue segmentation and structure segmentation. Tissue segmentation aims at classifying different types of brain tissues, mainly white matter, gray matter and cerebrospinal fluid (CSF). Although each tissue type displays distinct intensity in MR images, classification of voxels into a tissue type is difficult because of artifacts such as noise and partial volume averaging, where multiple tissues are present in a single voxel [Pham and Prince, 1999]. In order to compensate for these issues, the concept of fuzzy segmentation has been gaining interest, where voxels may be classified into multiple classes with a varying degree of membership. Fuzzy C-means (FCM) clustering has been employed by many researchers with different variations. [Pham and Prince, 1998; Pham and Prince, 1999] developed an adaptive FCM that can handle intensity inhomogeneity of MR images. Several methods are proposed to incorporate spatial neighborhood information into FCM, such as [Pham, 2001; Ahmed et al., 2002; Zhang and Chen, 2004; Shen et al., 2005; Forghani et al., 2007]. The above methods are fully automatic, which is desired and also practically feasible in tissue classification.

Structure segmentation, on the other hand, extracts the boundary of a particular anatomical structure. The difficulty and methodology of the segmentation highly depend on the structure to be segmented. Unlike

tissue segmentation, fully automatic methods are not always the best because certain amount of user control may be necessary to ensure the quality of segmentation. Since shape/size difference in individual brain structures between different populations is of our interest, we focus on structure segmentation in the rest of this thesis. A number of methods for general object segmentation can be applied to brain structure segmentation. They can be roughly divided into two categories: boundary-based approach and region-based approach. Boundary-based approaches, such as Snakes [Kass et al., 1988] and its variations, attempt to align an initial deformable boundary with the object boundary by minimizing an energy functional which quantifies the gradient features near the boundary. Due to their relatively reliable performance and low computational cost, they have been widely adopted in medical image segmentation. However, these methods can be easily trapped into local minima because of noisy edges, thus most of them require a good initialization near the solution or supervised by high-level guidance. Region-based approaches such as region growing [Hojjatoleslami and Kittler, 1995] assign membership to pixels according to homogeneity statistics. They appear to be more robust than boundary-based approaches since the information of the whole region is taken into consideration. However, since there is no easy way to distinguish boundaries and interior pixels of the object, these methods can lead to noisy boundaries and holes in the interior. Hybrid approaches [Ronfard 1994; Chakraborty and Duncan, 1995; Jones and Metaxas, 1997a; Jones and Metaxas, 1998; Chen and Metaxas, 2005; Jones and Metaxas, 1997b] combining the above two categories have become more and more popular recently since they can overcome the disadvantages of each type of methods alone.

Among boundary-based approaches, active contour (usually called deformable model in 3D) has achieved great success in medical image segmentation since its seminal work in [Kass et al., 1988]. In [Kass et al., 1988], the active contour is a parametric contour which is deformed according to image data with smoothness constraints. Direct implementation of this model cannot handle the topological changes of the evolving contour without explicit point manipulation. Usually, it can only segment a single object with simple topology. Some variations based on [Kass et al., 1988] can be found in [Radeva et al., 1995; Brigger and Unser, 1998; Heckenberg et al., 2006; Loizou et al., 2007]. Unlike the explicit parameterization of the active contour, level set implementation of the active contour [Osher and Sethian, 1988] represents the object boundary by a zero level set of a signed distance function, thus it allows for automatic changes in

4

topology. Novel geometric models of active contours have been proposed based on geometric flows [Caselles et al., 1993; Caselles et al., 1997; Malladi et al., 1993; Westin et al., 2000]. Recently, a fundamental work of active contour using Mumford-Shah function and level set implementation was proposed by [Chan and Vese, 2001]. This model can segment objects whose boundaries are not necessarily defined by image gradient, such as smooth boundaries. More related work about active contour models will be discussed in Chapter 2.

1.3. Quantitative Analysis of Brain Morphology

Quantitative morphologic assessment of individual brain structures is often based on volumetric measurements and shape analysis. Volume changes are intuitive features in explaining atrophy or dilation due to illness. However, shape analysis can precisely locate morphological changes in pathological structures which cannot be reflected in volume measurements, thus shape analysis has gained increasing interest from neuroimaging community.

Volume changes of brain structures have been extensively studied. For example, most previous studies found reduced size in the CC and its sub-regions in autism [Piven et al., 1997; Hardan et al., 2000; Hardan et al., 2009; Keary et al., 2009; Frazier et al., 2009]. Other examples of volume analysis on other brain structures can be found in [Hardan et al., 2007; Knickmeyer et al., 2008; Tsatsanis et al., 2003]. Besides volume measurements, some simple geometric features of brain structures have also been explored. For example, [Vidal et al., 2006] measured the thickness of the CC thickness at certain division borders. We consider both the volume and other geometric features as global traits. None of these traits provide sufficient details of shape information. In Chapter 4, we show some work about global traits analysis as part of our framework, but our focus is the shape analysis described below.

1.3.1 Shape Analysis

There are various methods of shape analysis with varying complexities. We divide shape analysis into three types: landmark analysis, pseudo-landmark analysis and whole surface analysis. Each type can be in either 2D or 3D, but the problems in 3D case are much more difficult.

Landmark analysis has been popular in neuroanatomical research because of its convenience and effectiveness in obtaining shape information. Landmarks are usually determined by anatomical

prominences of the biological structure of interest. Since landmarks of different individual shapes are inherently corresponded, they can be directly compared among different shapes. Various forms of landmark coordinates can be used to analyze the shape morphology. Euclidean Distance Matrix Analysis (EDMA) and Thin-plate spline (TPS) are two common ways to analyze the landmarks. EDMA uses landmark coordinate data to calculate all pairwise distances between landmarks [Burrows et al., 1999]. It is invariant to the coordinate system, which makes it biologically and statistically advantageous [Theodore and Richtsmeier et al., 1995]. For example, it was used to analyze the sexual differences in the inter-landmark distances of the CC [Ozdemir et al., 2007]. TPS [Bookstein, 1989] has been widely used to compare landmark configurations [Ozdemir et al., 2007; Bookstein et al., 2001; Fink and Zelditch, 1995; Lapeer and Prager, 2000; Rosas and Bastir, 2002]. The fundamental principle of TPS is the comparison of two different shapes by deforming one shape to the other. The expansion factors can be used as a measurement of the deformation [Bookstein, 1991]. The disadvantage of landmark analysis is that the surface at non-landmark location cannot be interrogated and thus important information might be lost in the analysis. Therefore, it is usually performed together with whole surface or pseudo-landmark analysis to reveal the shape morphology from different aspects.

Whole surface analysis represents the shape as discrete points on a surface in 3D case or a contour in 2D case. Shape analysis is usually conducted by comparing the coordinates of each point between groups of shapes using multivariate statistical analysis. This type of analysis can reveal the shape morphology at every surface point, thus giving more details than landmark analysis, but the point correspondence among different shapes must be established before any statistical analysis. Some examples of whole surface analysis in clinical applications can be found in [Wang et al., 2003; Ferrarinia et al., 2006; Styner et al., 2005; Csernansky et al., 2002; Dager and Wang, 2007; Oguz et al., 2008]. Note that no anatomical landmarks need to be identified in DMS analysis, so the point at the landmark position may not be exactly matched to the landmark position on another shape due to matching error. This is a disadvantage of whole surface analysis, because inaccurate matching of anatomical landmarks can cause more serious problem in the shape analysis than inaccurate matching of a random surface point.

Pseudo-landmark analysis lies between the above two types. Pseudo-landmarks are the points interpolated between anatomical landmarks. The interpolated points as well as the landmarks are analyzed in a way

6

similar to whole surface analysis. Pseudo-landmarks can be viewed as a subset of dense surface points. They can approximate the whole surface but are not as accurate as the original surface points. Pseudo-landmark analysis makes a compromise between surface accuracy and computational complexity, and it ensures the correspondence of anatomical landmarks. A set of anatomical landmarks need to be clearly defined, which limits the application of pseudo-landmark analysis because not all the biological structures have prominent landmarks. Some examples can be found in [Aguirre et al., 2007; Hennessy et al., 2005; Hammond et al., 2005; Hammond et al., 2007]. Most of these literatures focus on the human face because it has identifiable landmarks.

1.3.2 Point Correspondence in Shape Analysis

The main challenge in whole surface or pseudo-landmark analysis is the required step of establishing shape correspondence among individual shapes. In this way, each shape instance is represented by a certain number of corresponded landmark points, which facilitate the statistical comparison of shapes between different groups. Shape correspondence is a difficult problem especially in 3D case, because of the non-linear shape description and non-rigid shape variation.

Many methods have been developed for shape correspondence in recent years [Dalal et al., 2007; Davies et al., 2002; Styner et al., 2006; Dalal et al., 2009; Cates et al., 2006]. They can be classified into two types: mapping based and non-mapping based. Mapping based methods need to first map each shape instance onto a simple geometry. Spherical mapping is most commonly used in previous works, such as Minimum description length (MDL) [Davies et al., 2002] and spherical harmonics descriptors (SPHARM) [Styner et al., 2006]. In MDL, the shape correspondence error is measured by the required bit-length to transmit the resulting statistical shape model and all the training shape instances. SPHARM use first order ellipsoid from the spherical harmonic coefficients to align the spherical parameterization through which the correspondence among 3D shape instances is established. The spherical mapping can be applied to the analysis of a variety of shapes, but it is not appropriate for non-sphere-like shapes because of the extreme distortion in the mapping. Conformal mapping [Gu and Yau, 2003] has also been explored as a potential approach to the matching and analysis of shapes, which is more powerful and flexible in handling different shape topologies. [Gu et al., 2004] used spherical conformal map to match cortical surfaces, where they

parameterized the genus zero closed surfaces by spheres. However, mapping to other types of geometry can be used to handle non-sphere-like shapes as well, which will be explored in Chapter 3.

Non-mapping based approaches do not require any mapping of the shape surface and thus they are easier to understand conceptually. A simple method is to warp each shape to a template shape and for each point on the template the closest point on the target shape is the corresponding point. This method is usually used in pseudo-landmark analysis because the shape warping requires landmarks as control points. [Dalal et al., 2007] proposed a landmark sliding algorithm and the result is comparable to those of MDL and SPHARM. It first constructs an initial correspondence between two shapes by the simple method, and then refines the matched points on the target shape by sliding the points along the tangent direction. The sliding criterion is to minimize a shape correspondence error defined by TPS energy. [Dalal et al., 2009] further extended this method to open surfaces. [Cates et al., 2006] proposed an entropy based particle system where the correspondences are modelled as sets of dynamic particles, and the optimal configuration as one in which the entropy of each shape is balanced against the entropy of the ensemble of shapes. Applications of this method on brain structures can be found in [Cates et al., 2007; Oguz et al., 2008].

Mapping based methods are usually used for whole surface analysis since the mapping is implemented on the entire shape and no point constraint is needed theoretically. Non-mapping based methods are applicable to both whole surface and pseudo-landmark analyses, depending on the exact method used and the complexity of the shapes. Chapter 3 will focus on the methods for shape correspondence and the clinical applications are shown in Chapter 4.

1.4 Facial Morphology Analysis

Facial morphology analysis is an interdisciplinary subject that involves intensive knowledge of computer vision, statistics and clinical diagnosis. Most previous works focused on landmark analysis [Dalal and Phadk, 2007; Loos et al., 2003; Boehringer et al., 2006]. To our knowledge, no point-by-point shape comparison (i.e., pseudo-landmark analysis or whole-surface analysis) of any facial component has been conducted. A possible reason is that it is very difficult to manually delineate the shape of each facial component, as opposed to clicking a few landmark points on the face image. Shape information of some facial components can be an important clue in clinical diagnosis. For example, as mentioned in 1.1, MFP

group is diagnosed according to their distinct eye shape. Currently, the diagnosis of MFP is based on human observation, and no quantitative measurement is used to characterize their eye shape. We believe that the framework for brain morphology analysis introduced in 1.3 can be translated to facial morphology analysis. Therefore, robust and accurate facial feature localization, especially shape contour delineation, is in demand for the subsequent facial morphology analysis. Although there are previous works on 3D facial landmark analysis [Hammond et al., 2005; Hennessy et al., 2005; Mutsvangwa et al., 2007; Hennessy et al., 2006], we will only focus on 2D facial features since 2D contours are sufficient to characterize the shape feature of most facial components, such as the eye.

Currently, manual labeling of the facial landmarks is used in almost all clinical analysis despite many methods developed for automatic facial feature localization. The major concern in clinical analysis is the accuracy of the landmark location which puts the use of automatic methods at risk. Therefore, semiautomatic methods that allow manual post-processing are preferred. Even with manual post-processing, a semiautomatic method can still save a lot of time and labor compared to fully manual labeling, especially in outlining the shape of each facial component. A detailed review of existing methods on facial feature localization is presented in Chapter 5, where our new framework of facial feature localization is proposed as well. Both feature points (landmarks) and feature contours (shapes) can be located using our method.

After feature localization, the shape analysis framework can be applied to facial components the same as it is for brain structures. Most previous works treat the entire face as one organ, and analyze all the landmarks on the face. However, an individual facial component (such as the eye) may be of interest in some applications, and the shape of that facial component can be analyzed independent of other facial components. In Chapter 6, we will present the shape comparison of the eye among ASD, MFP and control groups. Our particular interest is whether the MFP group has a distinct eye shape compared to the rest of ASD population as observed by clinic diagnosis.

1.5 Organization and Contribution of the Dissertation

The three problems listed in 1.1 are addressed in Chapter 2~6. Chapter 2 presents several 2D and 3D segmentation methods for brain structures. Both automatic and semiautomatic methods are developed.

Chapter 3 proposes two methods for establishing 3D shape correspondence and chapter 4 gives several examples of statistical shape analysis of brain structures, where the methods in chapter 2 and 3 are applied. Chapter 5 proposes a new framework for facial feature localization on 2D images, and Chapter 6 gives an example of eye shape analysis.

Thus, the main contributions presented in this dissertation are as follows:

1. A context sensitive active contour for the segmentation of 2D corpus callosum is proposed. By dividing the active contour into several parts based on the prior knowledge about the corpus callosum shape, it can better overcome the local minima problem of the classic active contour.

2. A hybrid segmentation framework incorporating region based and boundary based methods is proposed. Region based approaches generate an initial seed that is as close to the object boundary as possible, making the subsequent boundary based approaches faster and more accurate. It can be used in both 2D and 3D with some variation in each case.

3. Two methods to establish point correspondence among 3D shapes are proposed. One is based on planar conformal mapping and the other is based on landmark sampling. Each method has its preferred applications in brain morphology analysis.

4. A new framework for automatic facial feature localization from 2D images is proposed. Both feature points and feature contours of the eyes, the nose, the mouth and the chin can be accurately located.

2. Brain MRI Segmentation

In this chapter we propose several methods for brain structure segmentation. Section 2.1 describes a semiautomatic method for 2D corpus callosum segmentation, section 2.2 describes a semiautomatic method for general 3D brain structure segmentation, and section 2.3 introduces a fully automatic framework for 2D brain structure segmentation. Section 2.4 gives a summary of this chapter.

2.1 A Context-Sensitive Active Contour for Semiautomatic 2D Segmentation

2.1.1 Background

Active contour models (snakes) have been widely used in image segmentation since their introduction [Kass 1988]. These methods iteratively update the contours according to various forces designed to seek out object/region boundaries while maintaining the smoothness of the contours. Interactively controlled forces may also be introduced to allow the user to guide the segmentation, which made active contours particularly popular for medical imaging applications. A survey of early work of deformable models in medical image segmentation can be found in [McInerney et al, 1996]. There is more recent work on snakes and their variants, such as [Brigger et al 1998, Wang and Chung, 2005]. The problem of most current active contour models is that they have little 'awareness' of where they are in the image, how their parts are arranged, or to what structures they are converging [McInerney et al, 2002]. Thus, there is a need to combine the low-level feature detection ability of active contour models with flexible high-level knowledge, which triggers the top-down and bottom-up combination scheme.

[Liang et al 2006] proposed a method called united snakes which combined snakes and live wire [Stalling et al, 1996] together. Live wire, or intelligent scissors, is an effective interactive boundary tracing tool. United snakes use live wire seed points as the hard constraints of snakes. This method can overcome a lot of disadvantages of a single live wire or snakes method. However, we have to carefully do a boundary tracing in order to get a sufficient number of seed points. [McInerney et al, 2006] proposed a sketch-initialized subdivision-curve Snake and applied it to 2D segmentation of the CC as well as other structures. This method is rapid, accurate and repeatable, but it requires a special pressure-sensitive device for user initialization. Recently, McInerney et al. introduced a novel approach for automatic medical image

11

segmentation that combines deformable model methodologies with concepts from the field of artificial life [McInerney et al, 2002]. They proposed deformable organisms that possess deformable bodies with distributed sensors, as well as brains with motor, perception, behavior, and cognition centers. Deformable organisms are perceptually aware of the image analysis process. Their behaviors, which manifest themselves in voluntary movement and alteration of body shape, are based upon sensed image features, pre-stored anatomical knowledge, and a deliberate cognitive plan. This method was further extended by [Hamarneh and McIntosh, 2005] which applied a physical-based implementation in order to provide an opportunity for the expert to intervene the segmentation intuitively. The results on CC segmentation in [Hamarneh and McIntosh, 2005] showed that minor user interactions could further improve the segmentation accuracy.

Inspired by the work of [McInerney et al, 2002], we introduce a new context sensitive active contour and apply it to 2D corpus callosum segmentation. Corpus callosum (CC) is the major communication pathway between the two cerebral hemispheres and mainly consists of axons. It is responsible for conduction of signals between homologous and heterotopic cortical regions and is an essential component for brain lateralization and interhemispheric communication. Structural changes, such as size and shape changes in the corpus callosum occur in a variety of neurological diseases, such as dyslexia, schizophrenia, autism, bipolar and unipolar disorder. Therefore, neurologists are interested in looking at the corpus callosum and analyzing its shape from MR images.

Although segmenting the CC seems simple, it turns out to be nontrivial. The global shape of the CC is relatively consistent, but the local shape variation is dramatic. The intensity of the CC also varies from one image to another. There can be gaps in the boundary of the CC almost anywhere, and parts of the CC may be narrow or have bumps. The most challenging problem is the existence of the fornix, which is a thin structure that may or may not contact the CC in the mid-saggital MR image. It is almost the same brightness as the CC, and the size and position of the contact region can vary considerably. Because of these challenges, traditional active contour models will not be robust enough to correctly extract the boundary in many cases. Manual tracing of CC boundary is still the most frequently used method in clinical studies, such as in [Christine et al, 2006, Piven et al, 1997, Palmer et al, 2002], which is time-consuming, error-prone and operator-dependent.

Unlike the fully automatic approach of [McInerney et al, 2002], our method is semi-automatic. It requires the user to interactively initialize a seed contour (through as little as three mouse clicks) that consists of four inter-connected parts. We divide the active contour into several parts according to our prior knowledge of the segmented object. Each part of the contour is assigned to a certain part of the object, and a set of deformation rules are designed for each part respectively. Each part of the seed is aware of its own orientation, its target structure, and more importantly can have its own motion law tailored for its corresponding target structure. By allowing the user to interactively initialize the model, our algorithm has a much lower computational complexity compared with the full automatic approach of [McInerney et al, 2002].

2.1.2 Contour Initialization and Partition

The anatomical structure of the CC is shown in Fig.2.1. We roughly divide the boundary curve of the CC into 4 parts (Fig.2.1), which are anterior (CA), posterior (BD), upper boundary (DC), and lower boundary (AB). Instead of using a closed-curve representation, we treat the four parts as four separate curves, which are connected by four sensor points. There are several advantages of using open-curve. First, it implicitly integrates our prior knowledge into the curve evolution process. In contrast, a closed curve can't be aware of any high-level information after it has been initialized. Secondly, this curve partition allows us to use different parameters of active contour evolution for each part, so that each curve can better fit the curvature feature of the boundary. Thirdly, it will make fornix detection easier, since we only need to search along curve AB to find the fornix.

The initialization requires the user to click three points to form a polyline $o_1 o o_2$ within the body of the CC (Fig.2.2(a)). After that, a seed contour is constructed in the following fashion. First, an edge map is generated through canny edge detection. Prior to edge detection and any other operations, we perform a Gaussian smoothing to the image in order to reduce the noise. The width σ of the Gaussian function is selected so that there is no small edge in the internal area of the CC. The edge map of the smoothed image is shown in Fig.2.3 (b).

We then do a point-tracing to find the four points on the edge (Fig.2.2(b)). Starting from o_1, we trace along a line perpendicular to line $o_1 o$ in both directions to find edge points a_1 and c_1. Each one is the first point on the edge along our tracing direction. Similarly, we trace from o_2 to find a_2 and c_2. In case there is a small

13

gap where the tracing line intersects with the edge, we may either find an edge point on the background, or never find an edge point until we reach the image boundary (the image boundary is treated as edge in our case). To avoid this problem, we trace along several lines whose angles with the original tracing line are within a small range [$-\theta, \theta$], where θ is a positive small angle. Among the first edge point on each tracing line, the one with the smallest distance to the starting point is selected. If the gap is so large that we can not find any edge point within [$-\theta, \theta$], we will increase the value of θ to cover a larger search range. However, in our experiment, θ=20° can satisfy all the images since there is usually high contrast on the upper and lower boundary of the CC.

We then show how to construct points b_1 and d_1 in Fig.2.2(c), and b_2 and d_2 follow the same way. We assume that there is always a small neighborhood around o_1 completely inside the CC region. This assumption is reasonable since the user can easily locate the end points in the middle with some margin to both the upper and lower boundaries. To best guarantee this assumption, we substitute o_1 with the middle point of line segment a_1c_1, denoted as o_1^*. Thus, we can draw a circle centered at o_1^* with a small enough radius (Fig.2.2(c)). b_1 and d_1 are two intersection points of line oo_1^* and the circle. Now we have the initial seed for each part of the curve—$a_1b_1c_1$ for anterior, $c_2b_2a_2$ for posterior, $a_2d_2od_1a_1$ for upper boundary and $c_1d_1od_2c_2$ for lower boundary. The four seeds are connected by four sensor points-- a_1, c_1, a_2, c_2. Fig.2.3 (a) and (c) show the initial polyline and the completed seed on the real image. Conceptually, this seed is a polygon with some edges overlapping in the body of the CC, while the four sensor points serve as both the separation of the entire curve and the connection of each curve segment. During the curve evolution process, these four points will not move since they are already on the boundary.

The performance of the curve evolution can be guaranteed only if the initial seed is completely inside the CC region. The seeds of left and right parts can be inside based on the above assumption. Normally, three points are sufficient to make an inside polyline. However, for some abnormal CC shapes more points may be needed, so in our algorithm we allow the user to click as many points as needed to make the initial polyline.

14
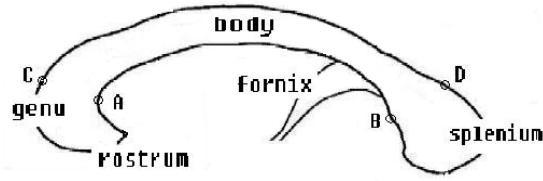
Fig.2.1. CC structure and partition of the CC boundary (image modified based on original illustration from [McInerney et al, 2002]).
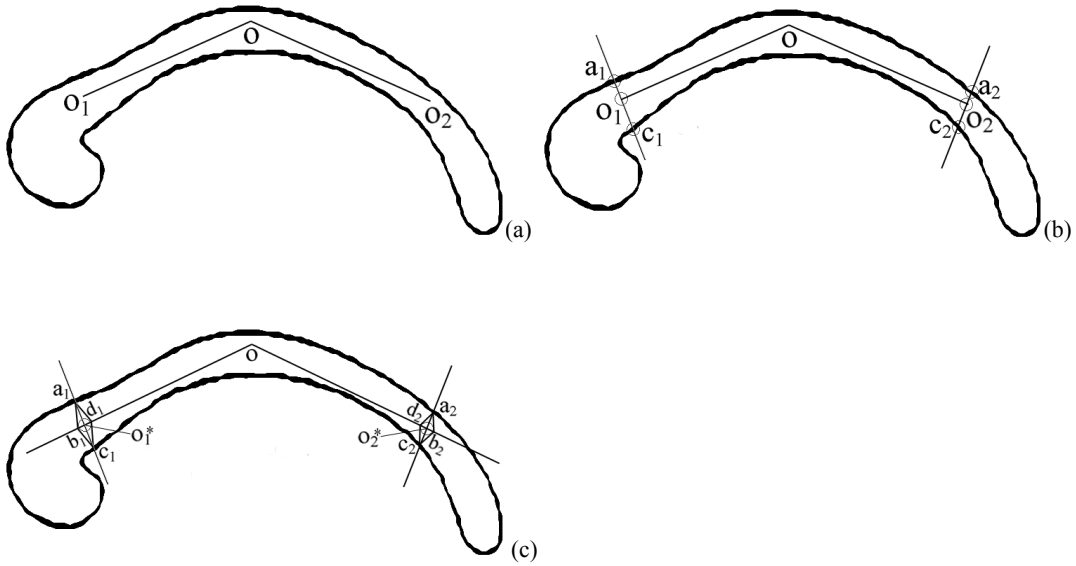
Fig. 2.2. Illustration of the seed initialization process. (a) User-initialized polyline. (b) Point tracing. (c) Completed seed contour.

Fig. 2.3. An example of the initialized seed. (a) User-initialized polyline. (b) The edge map. (c) The completed seed contour.

2.1.3 Contour Evolution

Consider a family of smooth curves $C(p, t)$ connecting two given end points, where p parameterizes the curve and t parameterizes the family. This family evolves according to the following partial differential equation [Belyaev 1999]:

$$\frac{\partial C(p,t)}{\partial t} = F\vec{n}, C(p,0) = C^0(p) \qquad (2.1)$$

15

where $\vec{n}$ is the unit normal vector of $C(p, t)$, F is the speed function, and t can be considered as the time parameter. The speed function F has the form similar to that in [Yushkevich et al., 2006]:

$$F = (v + \varepsilon k)g - \gamma(\nabla g \cdot \vec{n}) \qquad (2.2)$$

where k is the curvature of the curve, v is a positive constant speed, and ε, γ are two coefficients. g is a function derived from the input image I.

$$g(x) = \frac{1}{1 + \alpha(NG(x))^2}, NG(x) = \frac{\|\nabla(G_\sigma * I)\|}{\max_I \|\nabla(G_\sigma * I)\|} \qquad (2.3)$$

where $NG(x)$ is the normalized gradient magnitude at pixel x of image I, $G_\sigma * I$ is the image smoothed by a Gaussian kernel, ∇ is the gradient operator, and α is a coefficient. The first term in Equation (2.2) causes the curve to grow along its normal direction, and the second term acts in an opposite direction to the normal when the curve reaches the object boundary.

The orientation of the normal vector for each curve segment is specified upon initialization. We consider each curve segment as part of a closed curve drawn counterclockwise, and the normal orientation is defined as pointing outward of this closed curve. In order to get a correct normal orientation, the points on each curve should be arranged in a proper sequence.

We employ explicit Lagrangian approach to solve the curve evolution equation. Equation (2.1) can be approximated numerically by

$$C(p, t + \Delta t) = C(p, t) + (\Delta t)F\vec{n} \qquad (2.4)$$

where Δt is a small time step. Smooth curves are represented by polygonal lines composed of vertices $\{P_i\}_{i=1}^M$. In order to achieve extra stability and accelerate the convergence, a tangent speed component is added in the discrete curve evolution [Belyaev et al., 1999]:

$$P_i^{k+1} = P_i^k + \Delta t^k(F_i^k \vec{n}_i^k + G_i^k \vec{t}_i^k) \qquad (2.5)$$

where P_i^k indicates the location of vertex i in the kth iteration, $\vec{n}_i^k$ and $\vec{t}_i^k$ are the unit normal vector and tangent vector at P_i^k respectively, and F_i^k is a discrete approximation of F. The tangent speed component G is defined as

$$G_i^k = \frac{d_i^k - d_{i-1}^k}{d_i^k + d_{i-1}^k} \qquad (2.6)$$

where $d_i^k = P_i^k - P_{i+1}^k$ is the distance between two neighboring vertices P_i^k and P_{i+1}^k.

We use the method in [Duan et al., 2005] to estimate the time step Δt^k:

$$\Delta t^k = m_e / M_F^k \qquad (2.7)$$

where m_e is the unit grid cell length of the image data and M_F^k is the maximum magnitude of the speed F in

the kth iteration.

There are several ways to numerically calculate curvature and normal direction of each curve point. The

normal direction of a vertex is approximated by the average of the normal vectors of two incident line

segments. We use circular approximation in [Belyaev et al., 1999] to calculate curvature. Suppose *A, O, B*

are successive points on a curve, and we want to calculate the curvature of *O*. The circular approximation is

$$k = \frac{4S}{abc} \qquad (2.8)$$

where *a=|AO|*, *b=|BO|*, *c=|AB|*, and *S* is the area of triangle *ABO*. We assign a sign to the curvature of

each point. With the normal pointing outwards, the convex point has a negative curvature and the concave

point has a positive curvature.

All vertices are labeled active at the beginning. Vertex *i* terminates its motion when F_i^k is sufficiently

small [Belyaev et al., 1999], and it is then labeled as inactive. When all vertices are inactive, the curve

evolution stops.

To ensure that the numerical simulation of the curve evolution proceeds smoothly, we must maintain the

regularity of the curve. There are two issues to concern: point density and self collision. To maintain proper

density of the curve points, a new vertex is inserted between two adjacent vertices if the distance between

them is bigger than the maximum edge length, and a vertex is deleted if its distance with one of its

neighbors is smaller than the minimum edge length. The inserted vertex is initially marked active. The

proper edge length should be chosen so that the curve points are dense enough to capture the details of the

shape, while overhead points are avoided to reduce computational cost. In our case, the maximum edge

length is 4 and the minimum edge length is 1, with the length of the CC ranges from 130 to 160 (all in pixel

17

unit). To avoid self collision within each part of the curve, we apply a collision detection technique. Our method is similar to that in [Duan et al., 2005] while more efficient in convergence. If the distance between two non-adjacent vertices is smaller than the minimum edge length, we connect the two vertices and delete all the vertices in between.

2.1.4 Post-processing

The fornix may or may not contact the CC in mid-sagittal and surrounding slices. It is a thin structure with similar intensity level as the CC. For this reason, a standard active contour model can't extract the correct CC boundary connected to fornix. [McInerney et al., 2002] applied a find-fornix routine based on the parallelism between the lower boundary and the medial axis. [Lee et al., 2000] used a heuristic method to detect the feature points of the fornix. Compared with their methods, ours is much simpler. Since we know the fornix always appears beneath the body of the CC, we only need to search along the lower boundary to detect it. As shown in Fig.2.4(a), there are three distinct points around the fornix. If we find points a and b, we can connect them to get rid of the fornix tip. By carefully studying the curvature characteristics around the fornix, we design the following fornix removal algorithm. The flow chart of the algorithm is shown in Fig.2.5.

1. **Find c:** Search along the lower boundary curve segment to find the fornix tip (point c), which is the point with the smallest curvature.

2. **Check c:** if *curvature ratio* is bigger than a threshold, continue the next step; otherwise go back to **Find c**.

3. **Find a:** Search between point c and the left end of this curve to find the left corner of the fornix (point a), which has the biggest curvature.

4. **Check a:** If point a is active, go back to **Find c**. Otherwise continue the following steps.

5. **Find b:** Search between point c and the right end of the curve, and find point b such that the normal direction of point b is perpendicular to the line from point a to point b.

6. **Connect ab:** Connect point a and point b, resample points between a and b to maintain the curve regularity, and mark the vertices between a and b inactive.

In the above algorithm, the curvature is a signed value with the curvature sign defined in 2.1.3. In step 2, curvature ratio is the absolute value of the curvature of c divided by the absolute value of the average

curvature along the curve. In the case of Fig.2.4(a), this ratio is much bigger than one because the curvature of the fornix tip is conspicuously smaller than other points on the curve. However, there may be the case that the fornix is not connected to the CC, such as in Fig.2.4(b). Since there is no distinct fornix tip along this curve, 'Find c' routine in the initial algorithm may find an arbitrary point which happens to have the smallest curvature. In this case, the curvature ratio could not exceed a threshold due to the smoothness of the curve, and the algorithm will never do anything to the curve since it is always looped in 'Find c'.

Since the fornix tip will not appear until after a certain time of active contour evolution, if we started removing it too early, what we found may not be the real fornix. Step 4 is to make sure it is the proper time to remove the fornix. The resampling process in 'Connect ab' introduces new vertices between a and b, which would begin their motion toward the fornix tip again. We mark them inactive so that they will not form a new fornix after the end of the algorithm. The results with and without the fornix are shown in Fig.2.4(c) and (d).
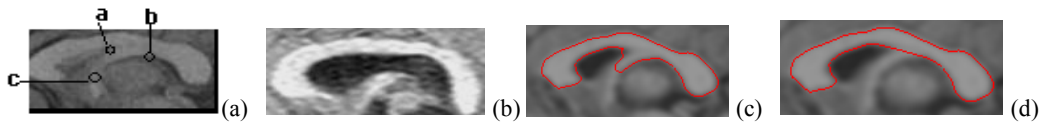
Fig.2.4. (a) Distinct points around the fornix. (b) The example of disconnected fornix. (c) The result without fornix removal. (d) The result with fornix removal.
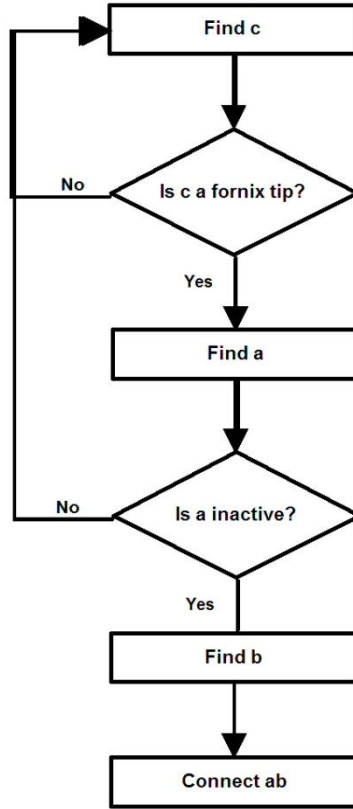
Fig.2.5. Flow chart of the fornix removal algorithm

2.1.5 Results and Validation

There are several parameters in equation (2.2) and (2.3). One advantage of our method is that we can use different parameters for different parts of the CC boundary. We find the four parts have different image features as well as curvature features. For example, the anterior and posterior curves have more high-curvature regions than the other two curves, and the upper boundary has more reliable edge than the lower boundary. Accordingly, ε should be larger for the middle parts and smaller for the left and right parts, and γ should be larger for the lower boundary and smaller for the upper one. In our experiment, the parameters in (2.2) and (2.3) fall into the following ranges: v=2~5, ε =0.5~2, γ =7~15, α =100~500, σ =2~3 for Gaussian smoothing of the image. The threshold of the curvature ratio in 3.6 is set to 10 in our experiment. We perform our algorithm on 2D brain MR images from different subjects. The experiment is running on an Intel(R) Pentium(R) D (2.8GHz) PC with Windows Vista. The total time for one MR image is 2-3 seconds.

Fig.2.6 shows the result on different MR slices of the same subject. The first column shows the fornix tip beginning to appear without fornix removal mechanism, the second column shows the results with the fornix removed, and the last column shows our result overlaid on the edge map. As we can see in Fig.6, the shapes of the CC look similar, but the image features at the boundaries as well as the fornix differ from one slice to another. Even if there is no explicit fornix on the image, there still might be a gap where the fornix is located, thus causing a dip of the lower boundary curve. As shown in Fig.2.6(c), our results can match the edge map accurately and bridge the gaps. Fig.2.7 shows the results on different subjects. The four columns are the same as in Fig.6, and each row represents the mid-sagittal slice of a different subject.

We use the measurements in [Yushkevich et al., 2006, Bueno et al., 2001, Chen and Metaxas, 2005] to quantitatively evaluate our segmentation results. We denote the correct segmentation result as C_{true}, our segmentation result as C_{seg}, and $| \bullet |$ as the area enclosed within the result. The following measurements are calculated.

1) False negative fraction (FNF), which indicates the fraction of structure included in the true segmentation but missed by our method:

$$FNF = | C_{true} - C_{seg} | / | C_{true} | \quad (2.9)$$

2) False positive fraction (FPF), which indicates the amount of structure falsely identified by our method as a fraction of true segmentation:

$$FPF = | C_{seg} - C_{true} | / | C_{true} | \quad (2.10)$$

3) True positive fraction (TPF), which indicates the fraction of the total amount of structure in the true segmentation that is overlapped with our method:

$$TPF = | C_{seg} \cap C_{true} | / | C_{true} | \quad (2.11)$$

4) Dice similarity:

$$Dice = 2 \times | C_{seg} \cap C_{true} | / (| C_{true} | + | C_{seg} |) \quad (2.12)$$

5) Overlap coefficient:

$$overlap = | C_{seg} \cap C_{true} | / | C_{true} \cup C_{seg} | \quad (2.13)$$

The last two measurements range from 0 to 1 with one indicating a perfect agreement between C_{true} and C_{seg}. The overlap measurement is a stronger test than Dice similarity for segmentation accuracy [Bueno et al., 2001].

We use the results of manual segmentation by a trained expert as the ground truth (C_{true}), and compare our results (C_{seg}) with the ground truth. The time for an experienced expert to segment the CC on one slice is about 20 minutes. The experiment is performed on MR images of 16 subjects, and the mean and standard deviation of each measurement across subjects are listed in Table 2.1. The table shows low FNF and FPF, and high values in the other three measurements. This result indicates our method has high accuracy.

Table 2.1. Quantitative validation results

	FNF	FPF	TPF	Dice	overlap
Mean	0.0689	0.0613	0.9525	0.9364	0.8803
Std	0.0613	0.0401	0.0438	0.0178	0.0304

(a)	(b)	(c)
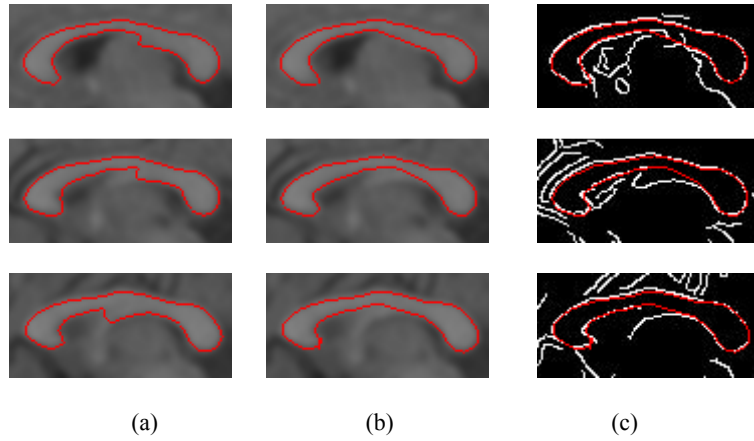
Fig. 2.6. (a) The fornix begins to appear without fornix removal. (b) The results with fornix removal. (b) The results overlaid on the edge map.

22

Fig. 2.7. (a) The fornix begins to appear without fornix removal. (b) The results with fornix removal. (c) The results overlaid on the edge map.

2.1.6 Conclusion and Discussion

We have designed some generic principles for the seed initialization. First, as stated in 2.1.2, the seed should be roughly within the body of the CC and not extend to the anterior and posterior parts. If the seed extends deep into the anterior or posterior part, the final curve will not be able to cover the whole anterior or posterior part. Second, the initial seed cannot have any portion outside the CC (Fig. 2.8 (a)). This is because the normal orientation of each curve is predefined given that they are all inside the CC. To ensure the seed is totally inside the CC, we allow the user to click more than three points if needed. Third, the length of the seed should cover the fornix gap, which means the right endpoint should be on right of the fornix (Fig. 2.8 (b)). This is to make sure that the fornix tip only appears on the lower boundary curve.

These principles are generally quite easy to follow. Since users can recognize the body of the CC as well as the fornix, they are able to locate the seed within the body and cover the fornix span without difficulty.

In conclusion, we propose a context-sensitive active contour method for 2D corpus callosum segmentation. This method takes advantage of prior knowledge by dividing the active contour into several parts and connecting them by sensor points. After a simple user initialization, a set of rules derived from prior knowledge will complete the initialization and guide the model deformation subsequently. The challenging problem caused by the fornix has been successfully solved. Experimental results demonstrate our method is fast, accurate and easy to operate.
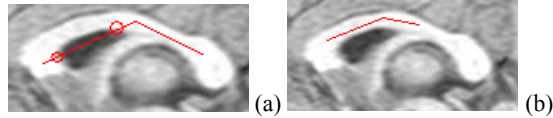
Fig. 2.8. (a) The portion between two circles is outside. (b) The seed not covering the fornix.

2.2. A Hybrid Method for Semiautomatic 3D Segmentation

2.2.1 Background

As discussed in 1.2, hybrid methods that combine region-based and boundary-based approaches are preferred to each approach alone. [Ronfard 1994] used region-based information to drive the explicit deformable models in their techniques, while [Chakraborty and Duncan, 1995; Jones and Metaxas, 1997a; Jones and Metaxas, 1998; Chen and Metaxas, 2005; Jones and Metaxas, 1997b] have addressed these issues by interlacing region-based and boundary-based methods into a united, iterative segmentation process. Especially, Jones and Metaxas [Jones and Metaxas, 1997a] proposed a hybrid approach based on fuzzy affinity clustering. Object boundaries are estimated based on fuzzy affinity metric of multiple image features. These estimations are recursively used to guide the deformable models and updated by the new model fit. They later extended the work in [Jones and Metaxas, 1997a] by an automatic initialization for multiple objects [Jones and Metaxas, 1998].

This section presents a hybrid method for 3D brain structure segmentation. It differs from previous hybrid methods in that it implements region-based and boundary-based approaches in two separate phases, which allows more efficient segmentation and effectively avoids local minimum. After a simple user initialization, a region-based technique is used to generate an initial seed contour that roughly represents the actual boundary, and a boundary-based method is used to guide the subsequent contour deformation. Since the seed generated by the region-based method is already close to the target boundary, the time needed for the subsequent boundary-based method is greatly reduced. The region-based method usually takes much less time than the boundary-based method, so our method is more efficient than a boundary-based method alone.

Level set implementation is preferred to Snakes especially in 3D segmentation because it avoids direct manipulation of the topology of triangle meshes. We use the level set model based on [Chan and Vese, 2001; Gibou and Fedkiw, 2005] as the active contour implementation. Moreover, we incorporate a

24

nonparametric kernel-based method [Heckenberg et al., 2006] into the segmentation of brain structures with low contrast boundaries, such as the thalamus.

Segmentation accuracy is crucial in clinical research. Hence, an expert is often needed to go through every segmentation result and make necessary adjustment. The most significant disadvantage of fully automatic methods is the lack of expert control over the segmentation, while semiautomatic methods have more flexibility by virtue of user intervention. We develop a user-friendly interface for user initialization and visualization. Moreover, we design a set of tools which allows the user to make modifications to the results after segmentation. In this way an expert can make the results more reliable and practical in clinical applications by post-segmentation adjustments.

2.2.2 Overview

The overall segmentation framework is shown in Fig. 2.9. The processes in round rectangles are operated by the user, and those in ellipses are automatic procedures. Depending on the intensity feature of the brain structure to be segmented, the user can choose to point initialization or region initialization. If the user initializes a point (solid boundary in Fig.2.9), the subsequent seed initialization and the deformation are performed on the original image. This is usually suitable for structures with high intensity contrast to the background, such as the corpus callosum. If the user initializes a region inside the brain structure (dashed boundary in Fig.2.9), a binary image is generated from a probability map [Heckenberg et al., 2006], and the subsequent procedures are based on the probability map which is updated during the deformation. This is useful for structures with similar intensity as the background such as the thalamus. Region based methods are used to generate the initial seed which is as close to the true boundary as possible, and boundary based methods (level set in our case) are used to refine the segmentation result by deformation. The details of the seed generation and the deformation will be discussed in the following sections. After the deformation, manual adjustment can be performed before or after the mesh extraction if the segmentation result is not perfect. The mesh extraction step converts the implicit volume representation to a triangle mesh for the ease of later analysis. A standard marching cubes algorithm [Lorensen and Cline, 1987] is applied here.

Fig.2.10 shows the user interface of the segmentation software we have developed. It is designed to be simple and easy to manipulate. It displays three orthogonal views of a volumetric image, along with a 3D view. The tools for user initialization and manual adjustment are on the right side.

25

Section 2.2.3 and 2.2.4 describe the point initialization and the subsequent contour deformation, section 2.2.5 describes the case of region initialization, and section 2.2.6 shows the experimental results.
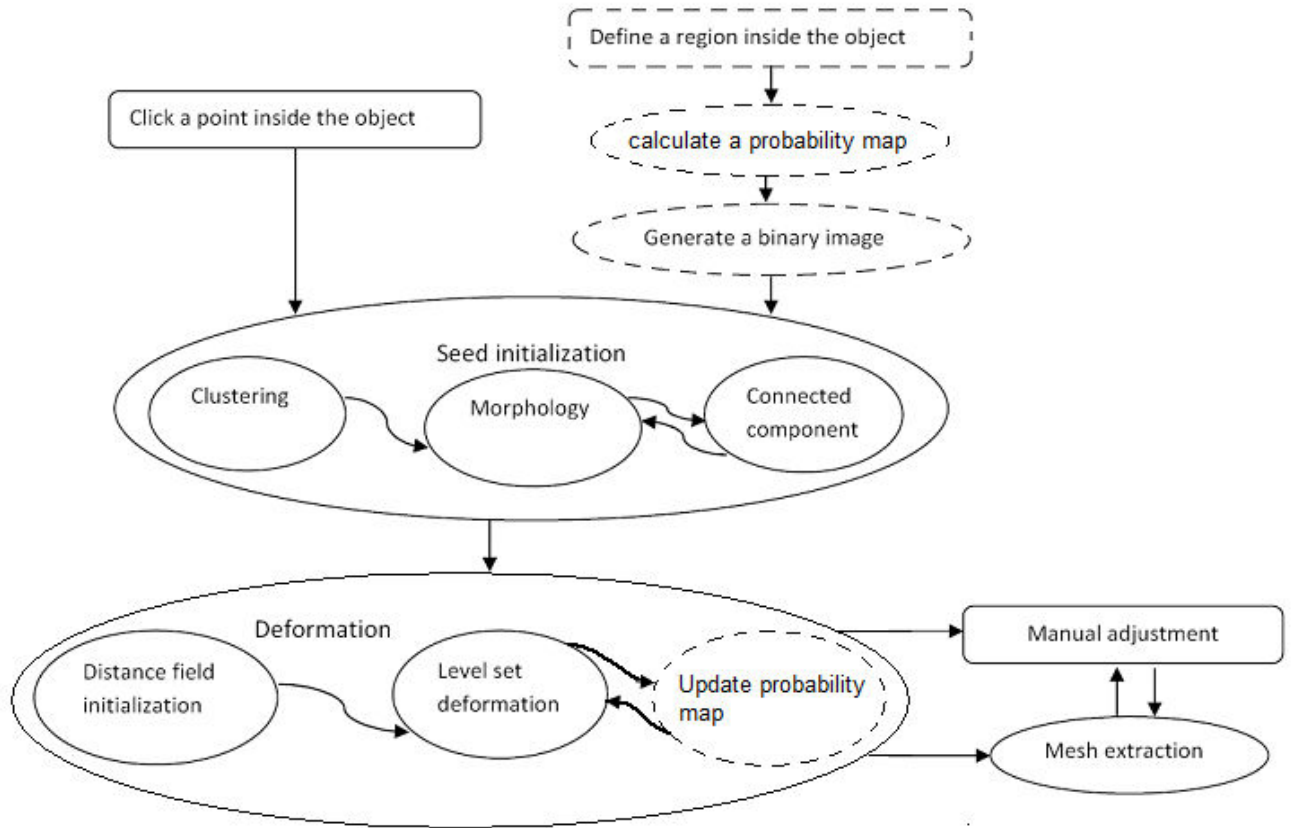
Fig.2.9 Flowchart of the segmentation procedure
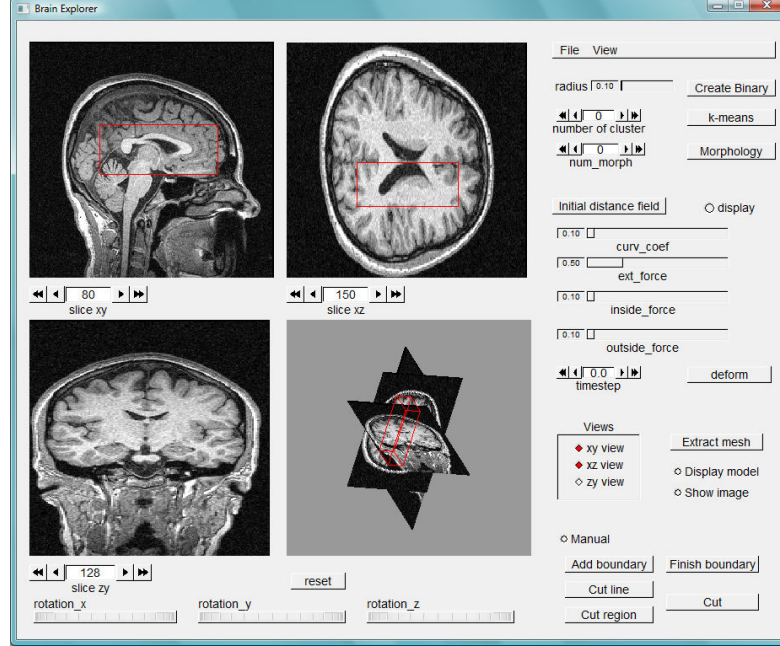
Fig.2.10 User interface of the segmentation software

2.2.3 Region Based Initialization

In order to improve the efficiency of the computation, the user can define a region of interest (ROI) before initializing the seed. An easy way to do this is to draw a rectangle on two of the orthogonal views and the two rectangles define a 3D ROI. Fig.2.10 shows the ROI of the ventricle defined on sagittal and coronal views. All the subsequent operations are performed within the ROI. If the ROI is not defined by the user, the entire image is the default ROI.

We use a region-based method to generate an initial seed which is close to the object boundary. The user is required to click a point $\vec{x}_0$ inside the target structure on one of the three orthogonal views. After that, clustering is performed based on the image intensity. We adopt fast adaptive K-means [Darken 1990] with a modified distance measure. Instead of L2 square distance, Delta-MSE dissimilarity [Xu 2004] is used as the distance measure. The Delta-MSE dissimilarity between data point $\vec{x}$ and the cluster centroid is defined as

$$D(I(\vec{x}), \bar{I}_j) = w_{ij} \parallel I(\vec{x}) - \bar{I}_j \parallel^2, \quad w_{ij} = \begin{cases} n_j/(n_j+1), p(\vec{x}) \neq j \\ n_j/(n_j-1), p(\vec{x}) = j \end{cases} \qquad (2.14)$$

27

where $I(\vec{x})$ is the image intensity at $\vec{x}$, $\bar{I}_j$ is the mean intensity of cluster j, $p(\vec{x})$ is the class label of

$\vec{x}$. The weight makes the dissimilarity bigger than L2 square distance if $\vec{x}_i$ is allocated in the cluster and

smaller otherwise. Therefore, the reassignments of data samples into their closest clusters are driven with

Delta-MSE dissimilarity more frequently than with L2 square distance, enabling a faster convergence to the

global optimum [Franti and Xu, 2004].

The cluster which contains $\vec{x}_0$ is the most important cluster because it provides an initial estimate of the

target shape. The number of clusters (k) is a parameter that the user can twist. However, in most cases $k=3$

or $k=4$ can guarantee a good estimate of the target shape. Let Ω_i denote the voxels contained in the cluster

that $\vec{x}_0$ belongs to and Ω_o the voxels outside the cluster. The next steps will further refine the seed (Ω_i)

by ensuring its connectedness and removing voxels from the seed which are not strongly affiliated with the

target structure. Voxels which are not affiliated with the target structure will in most cases be in or near the

boundary of Ω_i, denoted as $\partial\Omega_i$. To remove these voxels, Ω_i is eroded m times with a mathematical

morphology operation. This essentially removes $\partial\Omega_i$ from Ω_i at each iteration, and then recalculates Ω_i

after each step of erosion. In some cases, essential voxels of the target structure are removed from Ω_i, but

the deformation stage (see section 2.2.4) is used to overcome these seeding artifacts.

As with most mathematical morphology operations, Ω_i should then be dilated the same number of steps as

it was eroded. However, we insert a connected components step before dilation to ensure the connectedness

of Ω_i. It is possible that there are some other regions in the ROI that have similar intensity to the target

structure but are not connected to it, so they are likely to be in the same cluster with the target structure.

These artifacts can be removed by applying the connected components algorithm [Ronsen 1984]. The

algorithm is implemented originating from $\vec{x}_0$ so that the connected component containing $\vec{x}_0$ can be found.

During the connected components search, only voxels in Ω_i are available to be searched. Ω_i is

recalculated based on all voxels that are visited during the search, and the outside set is also recalculated

accordingly. Finally, the set Ω_i is dilated m times as it was eroded, recalculating Ω_i and Ω_o dynamically.

The value of *m* depends on the specific structure of interest, which will be discussed in our experimental results. These seed creation steps can be done very efficiently, and allow for the development of a seed that is roughly equivalent to the target structure. Fig.2.11 shows the seed initialization phase of a right lateral ventricle from start to finish.
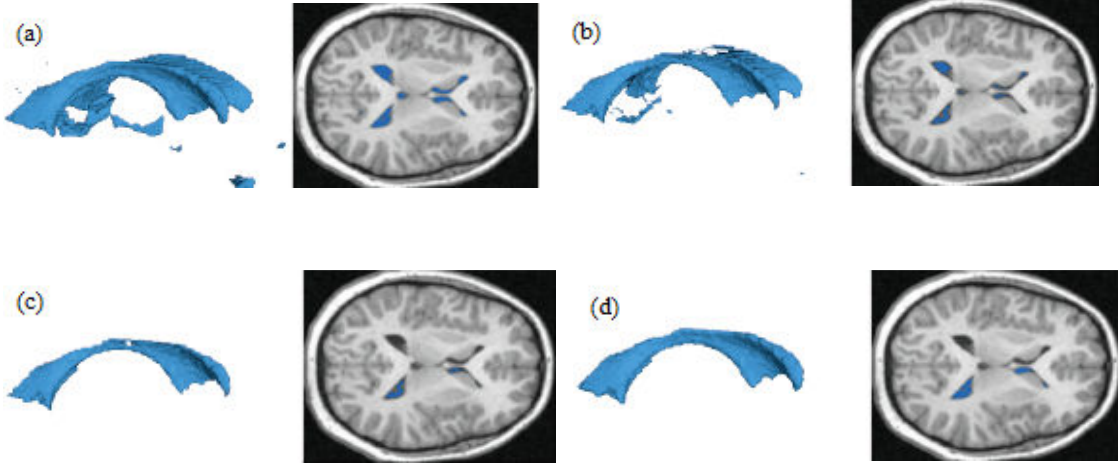
Fig.2.11 Demonstration of the seed generation phase of a right lateral ventricle. The 3D seed (left) is rendered explicitly for display purpose. On the 2D image slice (right), pixels are rendered as blue if they belong to the seed. (a) seed after K-means clustering; (b) seed after erosion; (c) seed after connected components search; (d) seed after dilation.

2.2.4 Boundary Based Deformation

The seed created in the previous section is coarse and contains various artifacts or imperfections due to image noise. This is the main drawback of region based methods in image segmentation. Therefore, the seed needs to be deformed under a boundary based mechanism in order to increase the accuracy of the segmentation. However, since this seed is already close to the target boundary, local maxima can be avoided in the following deformation, and the time needed for the deformation is greatly reduced compared to a point general point or sphere seed.

The seed Ω_i is transformed into an implicit function and deformed based on a level-set PDE. A signed distance function ϕ is initialized such that $\phi(\vec{x})$ is the signed Euclidean distance (negative on Ω_i and positive on Ω_o) from $\vec{x}$ to the closest voxel in $\partial\Omega_i$. This is implemented by a fast sweeping method

[Zhao 2005]. ϕ is initially very large at all voxels. The values for voxels in or adjacent to $\partial\Omega_i$ are directly computed with the following equation:

$$\phi(\vec{x}) = \begin{cases} \min_{\vec{y}\in\partial\Omega_i} \| \vec{x} - \vec{y} \|, \vec{x} \in \Omega_o \\ -\min_{\vec{y}\in\partial\Omega_i} \| \vec{x} - \vec{y} \|, \vec{x} \in \Omega_i \end{cases}$$

(2.15)

The values of the rest voxels are computed by propagating an estimation of the actual distance using previous computed elements. The details of the propagating algorithm can be found in [Zhao 2005].

ϕ is then deformed based on a PDE similar to those described in [Chan and Vese, 2001; Gibou and Fedkiw, 2005]. The PDEs in these works focus heavily on information other than edges to drive the level-set deformation. The idea behind the deformation is to introduce an artificial time variable t and to update the level-set function ϕ as time elapses. We update ϕ by numerically solving the following PDE (the variable $\vec{x}$ is omitted):

$$\partial\phi / \partial t = \{\alpha \cdot div(\nabla\phi / \| \nabla\phi \|) - \beta - \gamma_1(I - \overline{\Omega}_i)^2 + \gamma_2(I - \overline{\Omega}_o)^2\} | \nabla\phi |$$

(2.16)

where I is the voxel intensity, $\overline{\Omega}_i$ and $\overline{\Omega}_o$ are the average intensities of the voxels in the sets Ω_i and Ω_o respectively. α, β, γ_1, γ_2 are weight parameters. Specifically, α controls the curvature term, β is an external force to be applied to ϕ, γ_1, γ_2 are the weights for the distance functions $(I - \overline{\Omega}_i)^2$ and $(I - \overline{\Omega}_o)^2$ respectively. $| \nabla\phi |$ is set to 1 in [Gibou and Fedkiw, 2005]. The sets Ω_i and Ω_o are updated at each iteration to be consistent with their definitions, i.e. $\phi(\vec{x}) < 0$ for $\vec{x} \in \Omega_i$ and $\phi(\vec{x}) > 0$ for $\vec{x} \in \Omega_o$. For the purpose of efficiency, we also implement a narrow band algorithm [Lefohn et al., 2004] so that only values of the distance field that are within a certain threshold are updated. That is, we only solve equation (2.16) near the voxels where $\phi(\vec{x}) = 0$ when updating the distance field. The deformation will stop when an equilibrium is achieved. Fig.2.12 shows the results before and after deformation.

| (a) | (b) |
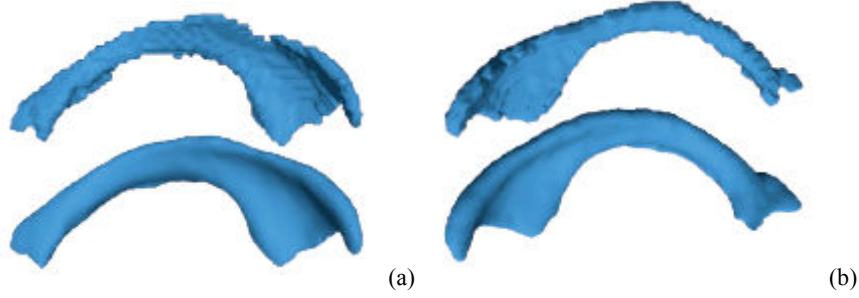
Fig.2.12 Seed contours rendered before (top) and after (bottom) level-set deformation. (a) right lateral ventricle (b) left lateral ventricle

2.2.5 Low Contrast Boundaries

In the case that a binary image is desired, the user can define a sphere M centered at $\vec{x}_0$ with the radius r (the circle on the 2D view in Fig.2.13(a)). The initial binary image is generated using a nonparametric method [Heckenberg et al., 2006]. Briefly, the intensity distribution is estimated based on Parzen-window function with Gaussian kernel. Suppose the volume of the initial sphere M is V, the probability of a voxel's intensity value $I(\vec{x})$ being consistent with the interior intensity inside the seed can be derived as

$$P(I(\vec{x})\,|\,M) = \frac{1}{V} \iiint_{\vec{y}\in M} \frac{1}{\sqrt{2\pi}\sigma} e^{-(I(\vec{x})-I(\vec{y}))^2/(2\sigma^2)} \qquad (2.17)$$

where σ is a constant that specifies the width of the Gaussian kernel, $\vec{y}$ is a voxel inside M. A threshold (usually the mean probability over the entire image domain) is applied to produce a binary image, where voxels with probability higher than the threshold have value 1 and other voxels have value 0 (Fig. 2.13(b)). The seed generation procedure is the same as in 2.2.4 except that it is performed on the binary image, so the seed before the deformation is generated based on the initial binary image in Fig. 2.13(b). In this case we can always let $k=2$ for the k-means clustering because there are only two intensity levels in the binary image. During the deformation the probability map is updated at each iteration according to (2.17) with M being the current model and V being the volume of M. The model deformation is performed based on the probability map instead of the original image. No binary thresholding of the probability map is needed during the model deformation phase. This method is especially suitable for brain structures with low contrast to the background. The difference between our method and the method in [Heckenberg et al., 2006] is that we apply region-based techniques on the binary image before the deformation, so that the seed is

31

close to the target boundary, while in [Heckenberg et al., 2006] a point seed is deformed from the beginning.

(a)   (b)   (c)
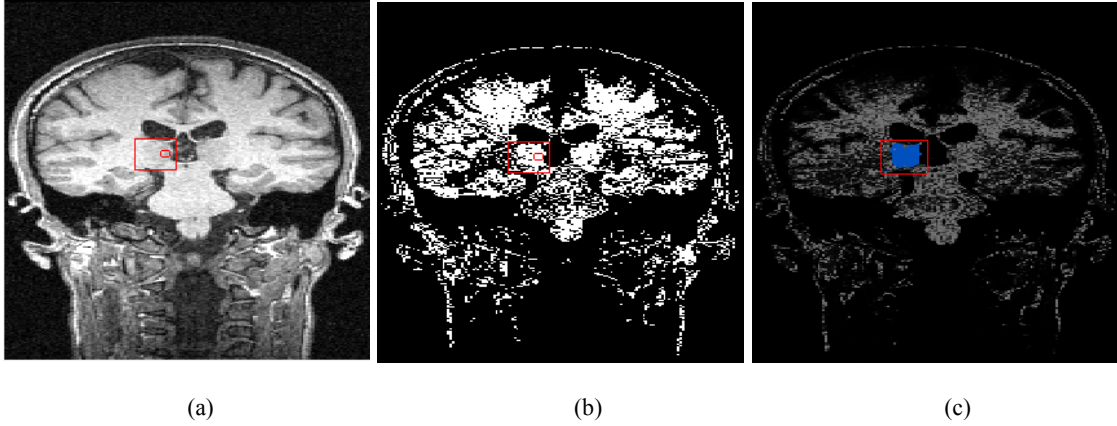
Fig.2.13 (a) initial seed of the thalamus defined by the user (circle on the 2D slice); (b) initial binary image generated from the seed in (a); (c) intermediate model during the deformation (blue region) overlaid on the probability map

2.2.6 Post-processing

In some extreme cases the above level set model cannot achieve a perfect segmentation, and manual adjustment is necessary to ensure the segmentation correctness. We design several convenient tools for manual adjustment after the automatic segmentation.

In the case that part of the brain structure is missing in the segmentation result (Fig.2.14(a)), the user can delineate the boundary of the missing part by sequential point clicking. The user-delineated boundary and the partial boundary of the original segmentation result form a closed region, and the voxels inside this region are moved to set Ω_i from set Ω_o. This is done by setting $\phi(\vec{x}) < 0$ for these voxels (Fig. 2.14 (b)).

If the result includes voxels outside of the structure of interest, two options can be performed. The user can draw a cutting line on a 2D slice (Fig. 2.14 (c)) and specify which side of the line is to be cut off. The voxels on this side of the line are moved to set Ω_o if they were in set Ω_i before (Fig. 2.14 (d)). The other option is to draw a rectangle that encloses the voxels that need to be cut off from the segmentation result (Fig. 2.14 (e),(f)).

All the manual operations have to be performed on a slice-by-slice basis, which might be time-consuming. However, the purpose of manual adjustment is to ensure the correctness of the segmentation result, and it is easier to make judgments on each 2D slice instead of a 3D model. Moreover, manual operation is not the

main part of our segmentation framework and there is limited amount of operation needed, so the time spent on manual adjustment is not a big issue. The experimental results will show the frequency of the data and the slices that need manual adjustment.
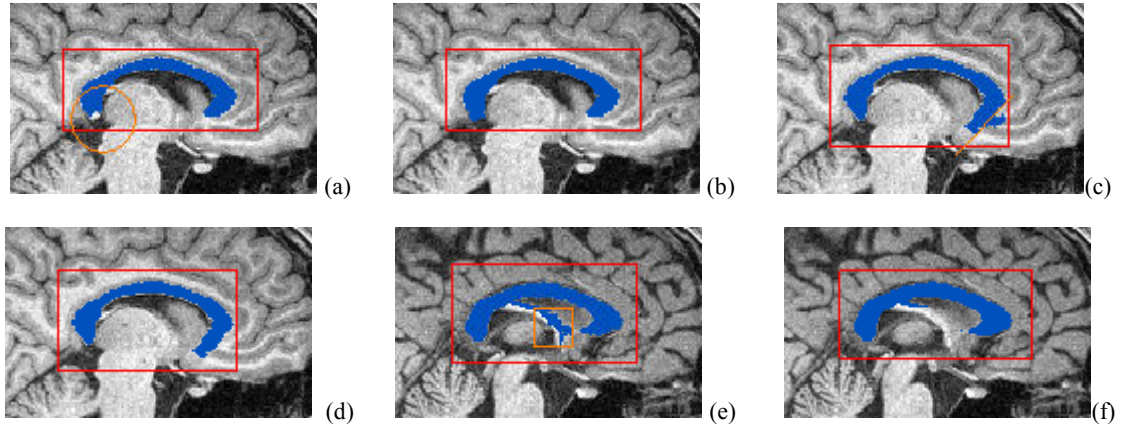
Fig.2.14 Demonstration of manual operations on the segmentation of the corpus callosum. (a) the splenium bottom is missing (the circle is used to highlight the missing part); (b) the result after manual modification on (a); (c) the segmentation result extends the anterior boundary and the user specified cut line; (d) the result after cutting off the region on right side of the line in (c); (e) the fornix enclosed in the small rectangle is to be cut off; (f) the result after cutting off the fornix in (e)

2.2.7 Results and Validation

The segmentation algorithm is implemented in C++ and the user interface is developed using FLTK. We test our method on two datasets. The first dataset consists of brain MRI scans of 20 college student volunteers obtained on a 3T Siemens Trio scanner with a standard 8-channel head coil in Department of Psychological Sciences at University of Missouri (TR=1920 ms, TE = 3.75 ms, flip angle = 8°, in-plane resolution = 1 x 1 mm, slice thickness = 1 mm, number of images = 160, matrix =256 x 256). The second data set consists of 25 children with autism recruited from the Thompson Center for Autism and Neurodevelopmental Disorders. T1-weighted brain MR images were acquired using the Siemens Symphony 1.5 T scanner (TR = 35ms, flip-angle = 30 degrees, thickness = 1.5mm, matrix =512x512).

Three structures are segmented using our method: the corpus callosum (CC), the lateral ventricles and the thalami. Fig.2.15 shows the segmentation results (after manual adjustment if necessary) in 2D and 3D views of 3 randomly selected subjects from each dataset. For each structure, one orthogonal view that best displays the features of this structure is shown, i.e., sagittal view for the CC, coronal view for the ventricles

and axial view for the thalami. The segmentation result is overlaid on the corresponding orthogonal view. A 3D view of all segmented structures is also shown. For the CC and the ventricles we choose the initialization method described in 2.2.3, and for the thalami we choose the initialization method described in 2.2.5. The deformation parameters are set in the following range: α =0.02~0.06, β =0.02~0.05, $\gamma_1 = \gamma_2$ =0.1, time-step=0.2. The seed generation parameters depend on the intensity distribution of the segmented structure and the background. We find that the best parameter values for the CC are: k=3, m=1, for the ventricles: k=3, m=0, and for the thalami: r=1.9~2.2, k=2, m=1.

With the facility of manual adjustment, our segmentation software can achieve both efficiency and optimal reliability, which is desired in clinical applications. Manual adjustment is made by a knowledgeable expert using our manual tools in 2.2.6 if necessary. From our 45 datasets, no manual adjustment is needed for the ventricle segmentation, 10 datasets need manual adjustment for the CC segmentation and 12 need manual adjustment for the thalamus segmentation. The slices that need to be modified usually appear at the end of the structure in its dominant view (eg., sagittal view for the CC). Among the data that need manual adjustment, the average number of slices adjusted is 6 for the CC and 4 for the thalamus.

To test the accuracy and efficiency of the segmentations, we compare our method with a generic boundary-based method (which uses a sphere seed centered at the user-specified point). The only difference between these two methods is the way that the seed is created. The segmentation results of the two methods are both compared to manual (ground truth) segmentations. The manual segmentation is performed by a trained expert different from the one who does the manual adjustment. For the comparisons, we calculate Dice similarity (equation 2.12) and overlap coefficients (equation 2.13) between manual segmentations and the results of each method. The range of these statistics lies between 0 and 1, with 1 indicating a perfect agreement between our segmentation and the manual segmentation. Table 2.2 provides a summary of the results. Overall, both methods are able to closely achieve the accuracy of the manual segmentations, but our method has higher accuracy than the generic boundary-based method. Manual segmentation takes 2 hours for each individual structure on average. To segment one structure using our method, the user initialization takes 20 seconds, the automatic algorithm part takes 90 seconds and the manual adjustment takes 10 minutes on average. Therefore, our segmentation framework is much more efficient than manual segmentation, even when the manual adjustment is needed.

34

We also show the efficiency of our segmentation method by comparing the running time of the automatic segmentation part (after user initialization and before manual adjustment) between the two methods. The results are shown in Table 2.3, which indicates our method is much faster in the automatic part than the generic method. Given the previous comparison of the accuracy, it is demonstrated that our new seeding scheme increases both accuracy and efficiency.

Since our method is not fully automatic, human factors may affect the consistency of the segmentation results. To test the effect of human factors, we have two operators perform the segmentation and manual adjustment on the same subset of the data. 20 datasets are randomly selected as the subset of the data. The two operators have been trained enough to be able to perform manual adjustment properly. Each operator is blind to the seed initialization and ROI definition of the other. The overlap of each pair of the segmentation results is measured by Dice coefficient. Table 2.4 shows the statistics of the Dice coefficients on different segmented structures. The high Dice similarity shows that our method can produce consistent results despite the user initialization.
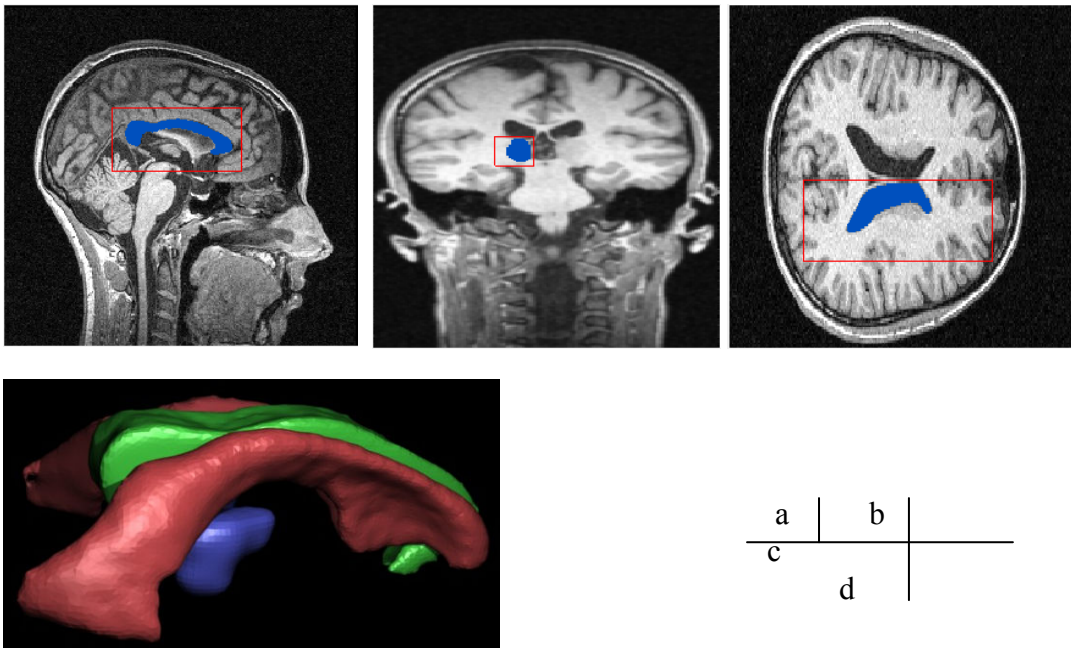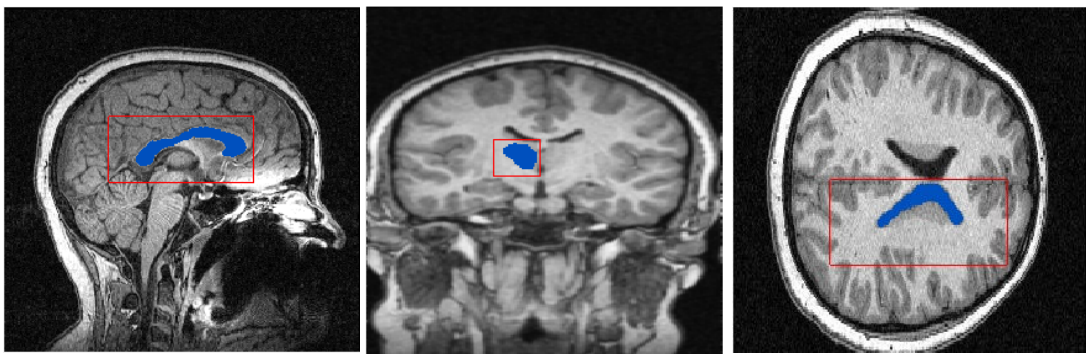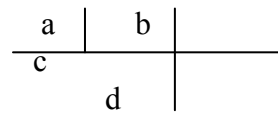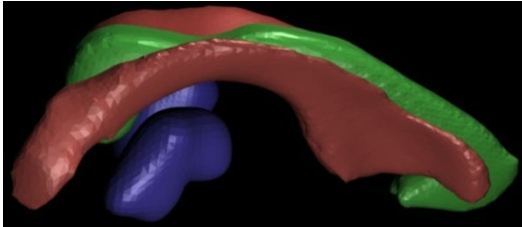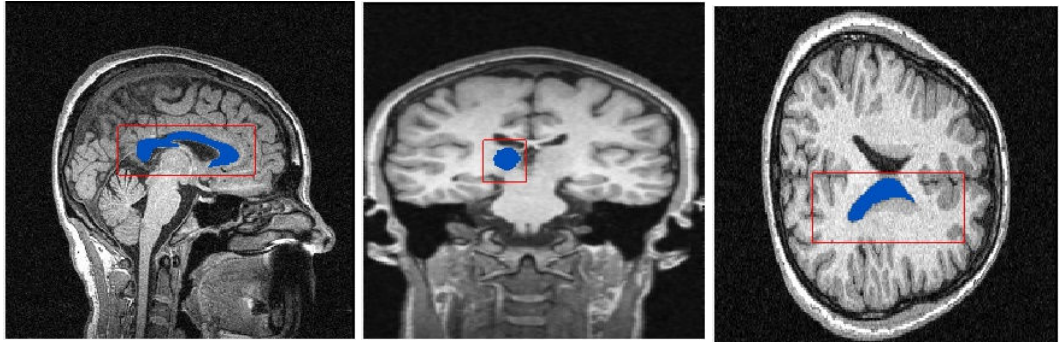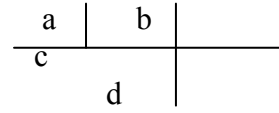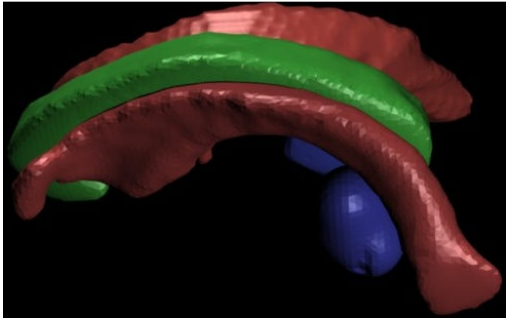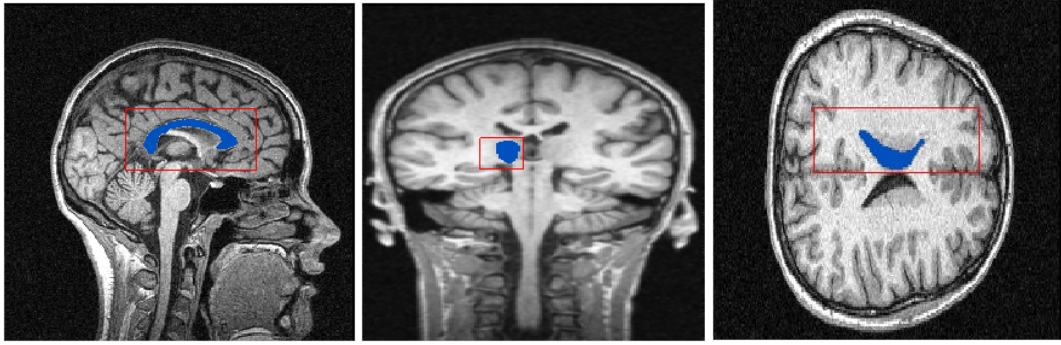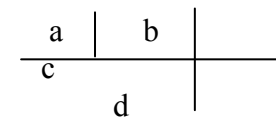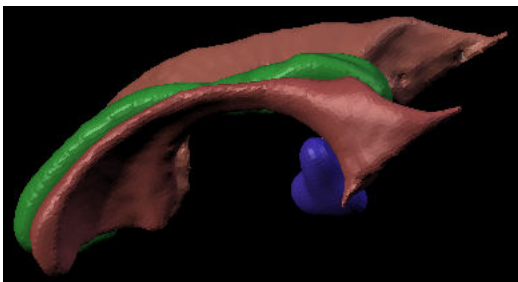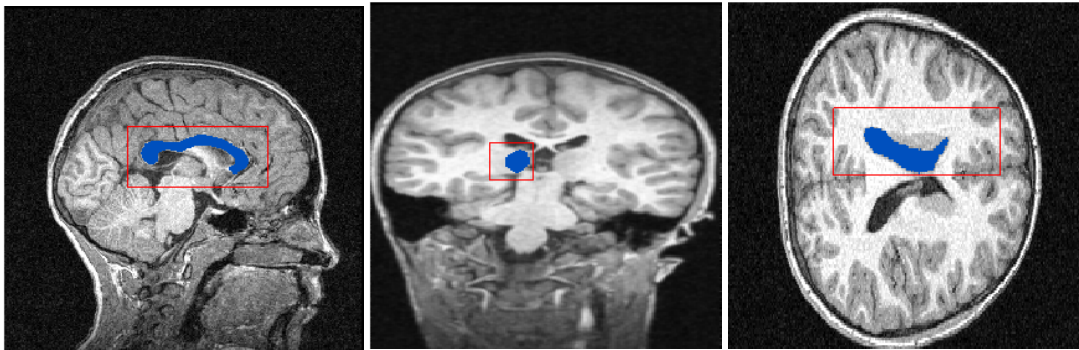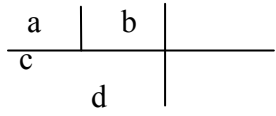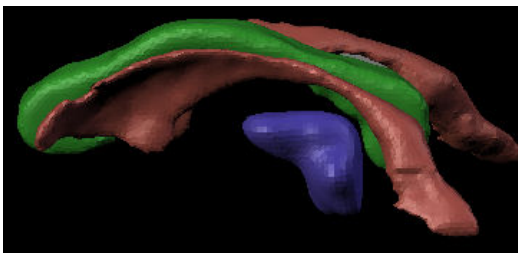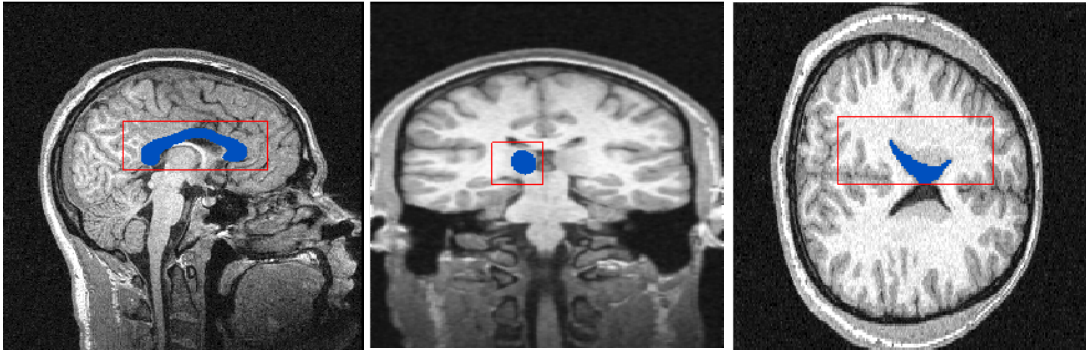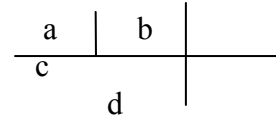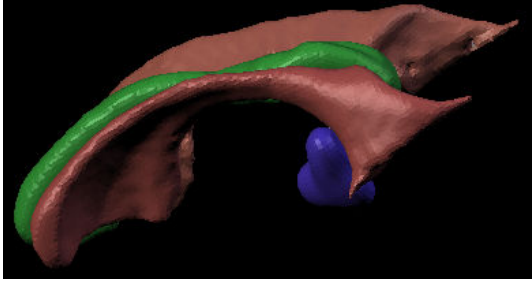
Fig.2.15 Segmentation results of student volunteers (the first three) and autistic children (the last three). (a) sagittal view of the segmentation result of the CC (blue region) (b) axial view of the segmentation result of the thalamus (c) coronal view of the segmentation result of the ventricle (d) 3D view of the segmented structures

Table 2.2 Accuracy measurements of our results and the results from the generic boundary-based method

Structure	Dice				Overlap			
	Our method		Generic method		Our method		Generic method	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CC	0.79	0.05	0.76	0.05	0.79	0.03	0.75	0.03
Left ventricle	0.81	0.04	0.78	0.05	0.73	0.05	0.72	0.05
Right ventricle	0.78	0.06	0.75	0.06	0.81	0.03	0.78	0.03
Left thalamus	0.69	0.08	0.64	0.09	0.66	0.06	0.63	0.07
Right thalamus	0.70	0.07	0.67	0.09	0.65	0.05	0.64	0.07

Table 2.3 Automatic segmentation time of our method and the generic boundary-based method (in seconds)

Structure	Our method	Generic boundary based method
CC	98	240
Left ventricle	75	182
Right ventricle	78	186
Left thalamus	88	124
Right thalamus	84	120

Table 2.4 Dice coefficients of the segmentation results from two raters

Structure	Mean	Std
CC	0.89	0.02
Left ventricle	0.94	0.01
Right ventricle	0.90	0.01
Left thalamus	0.82	0.07
Right thalamus	0.85	0.08

2.2.8 Conclusion

We present a semi-automatic 3D segmentation framework for brain MRI which combines region-based and boundary-based approaches. After user initialization, a region-based technique is used to generate an initial seed contour that roughly represents the actual boundary, and a boundary-based method is used to guide the subsequent contour deformation. In order to deal with brain structures with low intensity contrast to the

background, we choose to generate a binary image iteratively based on the intensity distribution of the deformable model, thus facilitating the segmentation. A user friendly interface is developed for user initialization and visualization. Moreover, a set of tools are designed that allows the user to make modifications to the results after segmentation. The results show that our method has similar accuracy to the manual segmentation but is much faster even with manual adjustment after the segmentation.

Our segmentation framework is a practical tool in clinical research. Our method can overcome the unreliability of fully automated methods in that we allow manual adjustment to the results, and it is much faster than pure manual segmentation. The user initialization is also very simple for a general user.

2.3. A Hybrid Method for Automatic 2D Segmentation

2.3.1 Overview

The segmentation methods described in 2.1 and 2.2 all require user initialization to some extent. In this section, we propose a fully automatic method for 2D MRI segmentation. Similar to 2.2, it is a hybrid approach that combines region based and boundary based approaches, and the prior information of the object to be segmented is utilized. However, the initialization is achieved by a set of recognition schemes so no user interaction is needed. We apply our method to the segmentation of the CC and the thalamus. Based on the shape and intensity characteristics of these two structures, we design two different inference schemes for the model initialization, but the segmentation framework is the same which can be summarized as follows.

1. A k-means clustering is performed based on the image intensity.
2. A hierarchical inference scheme is used to recognize the structure to be segmented, and an initial seed is constructed from the recognized region.
3. The seed is evolved under certain deformable model mechanism.

Fig.2.16 shows the segmentation procedures of the CC and the thalamus. The contour deformation mechanism is similar to 2.1 and 2.2, so we only briefly describe the deformable models (2.3.2) and then focus on the initialization in the rest of this section.

2.3.2 Model Deformation

39

Different deformable models are used for the CC and the thalamus. As in 2.1, the Snake model is used for the CC. This is mainly because the explicit representation of the boundary contour suits the fornix detection task better. The existence of the fornix often makes a single active contour fail if no user intervention is involved, because it is almost the same brightness as the CC. Thus an automatic fornix detection and removal needs to be performed after the segmentation. Equations (2.1)~(2.3) define the contour evolution mechanism, and we restate them below for convenience.

$$\partial C(p,t)/\partial t = F\vec{n}, C(p,0) = C^0(p)$$
$$F = (v + \varepsilon k)g - \gamma(\nabla g \bullet \vec{n}) \tag{2.18}$$
$$g(x) = \frac{1}{1 + \alpha(NG(x))^2}, NG(x) = \frac{\| \nabla(G_\sigma * I) \|}{\max_I \| \nabla(G_\sigma * I) \|}$$

The details and the numerical implementation of (2.18) can be found in 2.1. The parameters in (2.18) need to be carefully selected in order to obtain the best performance of snakes. However, we find that one set of parameters can work for all MRI data acquired under the same condition. In our experiment, $v=2$, $\varepsilon =0.2$, $\gamma =0.1$ and $\alpha = 10$.

On the contrary, the thalamus on the axial view usually has a much simpler shape compared to the CC shape, but the boundary is not as clearly defined by gradient as in the CC shape. Therefore, we employ the level set method [Chan 2001] as in 2.2. The contour evolution function is defined in (2.16), which is restated below.

$$\partial \phi / \partial t = \{\alpha \cdot div(\nabla \phi / \| \nabla \phi \|) - \beta - \gamma_1 (I - \overline{\Omega_i})^2 + \gamma_2 (I - \overline{\Omega_o})^2\} | \nabla \phi | \tag{2.19}$$

This formula allows the level set to converge using the pixel intensity inside and outside of the level set instead of gradient, which is well suited for the case of the thalamus boundary. We set $\mu = 0.5, \gamma_1 = 1, \gamma_2 = 1$ in our experiment.

2.3.3 Hierarchical Inference Scheme

We perform a k-means clustering based on the pixel intensity. Other clustering methods can also be used and we choose k-means because it is simple. In order to find the initial region for the deformable models from the clustering results, we design a hierarchical inference scheme which searches the interested object from coarse level to fine level, and then an initial seed for the deformable model is generated from the detected sub-cluster. Since CC and thalamus are quite different in intensity and shape features, direct and

indirect schemes are designed to recognize CC and thalamus respectively, but they both follow the same pipeline.

In the coarse-level inference, the goal is to find a cluster of interest from all clusters generated by k-means. The number of clusters K needs to be selected so that most of the pixels that belong to the CC are grouped into the same cluster. From our experiment K ranging from 4~10 can give a reasonable result for most cases, and we find K=8 generates the best result. Since the intensity of the CC is higher than most other parts of the image, we can leave out the clusters with relatively low intensity. In order to take into account the image noise which can have higher intensity than the CC, we first select the top three clusters with highest intensities, and then select the one with the largest area (number of pixels) as the CC cluster, because we observe that the noise regions are usually very small compared with the CC. Fig.2.17(a) shows the selected cluster which contains the CC.

For the thalamus, it's difficult to directly recognize the cluster that contains it since the intensity of the thalamus is not distinguishable from other parts of the image. Therefore, a different scheme is used for the coarse level inference. We notice that the ventricle (Fig.2.16(e)) is always on top of each thalamus, which has the lowest intensity in the image. We come up with an indirect inference scheme which infers the position of each thalamus from each ventricle. Only two clusters are needed in the k-means clustering since the intensity of the ventricles is very low, and there are no other dark regions connected to the ventricles. The cluster recognition is thus reduced to finding the cluster with the lower intensity.

The selected clusters (black areas) in Fig.2.17 (a) and (b) contain many false positives, i.e., other parts of the image which are not the object of interest. In the fine-level inference, the goal is to remove those false positives from the selected cluster. We first split the cluster into numerous sub-clusters by connected components algorithm [Ronsen 1984], so that each sub-cluster is a connected component. In the case of the CC, the object of interest is the sub-cluster that contains the CC, and all the other sub-clusters are false positives. Fuzzy set theory [Zadeh 1965] is used to detect the CC sub-cluster. We extract three features of the CC which are the area, length and width. The length and width are measured on the oriented bounding rectangle, and the area is the number of pixels that constitute the CC. Each feature is associated with a fuzzy set. A fuzzy set A is characterized by its membership function $\mu_A(x):[-\infty,+\infty] \to [0,1]$. The value of $\mu_A(x)$ indicates the degree to which the feature value x belongs to the fuzzy set A. A trapezoid

41

membership function is constructed for each fuzzy set, and the parameters are the center, the length of the top base (short) and the length of the bottom base (long) of the trapezoid. In order to accurately define each membership function, we randomly select a small number of images to learn the parameters prior to the segmentation. All the images are sagittal MR slices on or close to the mid-sagittal plane, and the pixel size is normalized to 1mm. The boundary of the CC is manually delineated on those images. If the fornix is present on the slice, the boundary will include the fornix as well because the fornix is usually in the same sub-cluster as the CC is in (Fig.2.17(a)). Then the three features of the CC or CC-fornix conjunction are calculated for each image. For each feature, the center of its trapezoid membership function is the average value of this feature, the length of the top base is twice the standard deviation of this feature, and the length of the bottom base is five times the standard deviation. The membership function can serve as a scoring function in decision making. For each sub-cluster, its area, length and width are calculated and applied to the corresponding membership functions defined above. Each membership function gives a score of this sub-cluster based on a single measurement. An overall score is computed as the weighted summation of the three single scores. The weights can be set according to the impact of each feature, but we find that equal weights for the three scores can achieve a good result. The sub-cluster with the highest overall score is selected as the CC sub-cluster. The detected CC sub-cluster is shown in Fig. 2.16(b).

In the case of the thalamus, the false positive of the ventricle cluster (Fig.2.17(b)) needs to be removed. We notice that the two ventricles are almost symmetric, so we detect them together based on the symmetry. Five features are measured for each sub-cluster, including the three features used to describe the CC sub-cluster. The other two features are the absolute angle with the vertical line and the y coordinate of the sub-cluster center. Fig.2.17(b) shows these two features of one ventricle sub-cluster. We then consider all possible pairs of the sub-clusters. For each pair the similarity between each feature value of the two sub-clusters is calculated as the ratio of the smaller value to the larger value, so that it lies in the range [0,1]. Similar to the overall score for the detecting the CC sub-cluster, an overall similarity is the weighted summation of the five ratios. We find that for a pair of ventricles there can be larger difference in their shape and size than in their orientation and altitude, so we set higher weights for the angle ratio and the center ratio, and lower weights for the other three. The pair with the highest overall similarity is selected as the pair of ventricle sub-clusters. The detected ventricle sub-clusters are shown in Fig. 2.16(f).
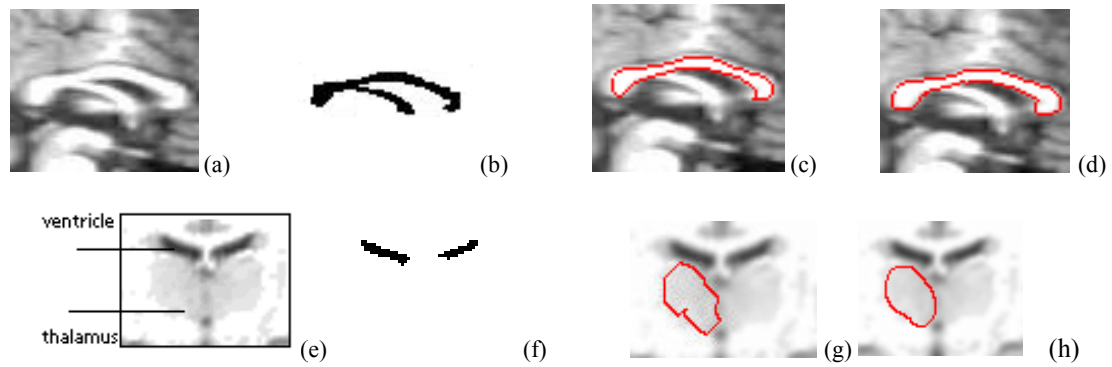
Fig. 2.16. Segmentation procedure of the CC (a-d) and the thalamus (e-h): (a,e) part of an original image (b,f) recognition result (c,g) initial seed (d,h) final contour
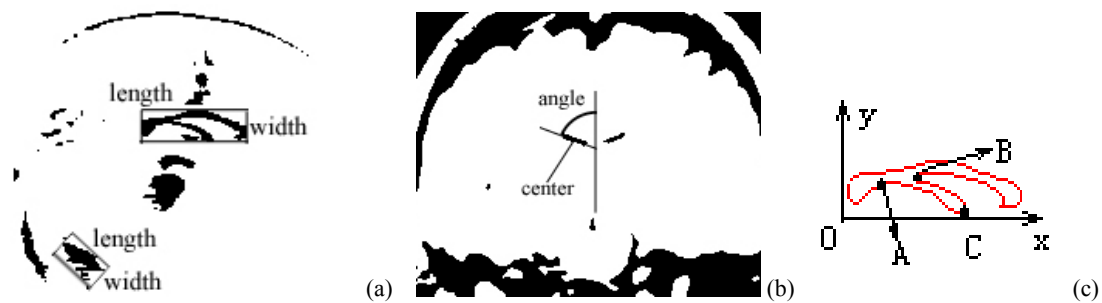
Fig. 2.17. (a) the CC cluster (black) and the oriented bounding rectangles of the CC and a noise region (b) the ventricle cluster (black) and the additional measurements of one ventricle (c) three distinct points on the fornix (A,B,C)

2.3.4 Automatic Seed Generation

After the detection of the sub-cluster, an initial seed can be easily constructed for the subsequent model deformation. Different methods are used for the CC and the thalamus. Especially, the fornix in the CC sub-cluster needs to be removed before the seed construction.

Since explicit active contour is to be used for CC segmentation, a boundary contour is extracted from the recognized CC sub-cluster using 2D marching cube [Lorensen and Cline, 1987], and a Laplacian smoothing is performed on the contour. Before contour extraction, we first dilate the region a certain number of steps based on the image size and then erode it the same number of steps using standard mathematical morphology in order to fill the holes inside the region. The extracted CC contour may contain the fornix, because in most cases the fornix is grouped into the same connected component as the CC. Similar to section 2.1.4, we try to detect the two points connecting the fornix and the CC (A, B in Fig.2.17(c)) and then connect them to remove the fornix, but a different detection method is applied here.

43

We define three landmarks of the fornix, which are points A, B, C in Fig.2.17(c). A and B are the points connecting the CC and the fornix, and C is the fornix tip. We use a method similar to the inference scheme for the CC sub-cluster detection in 2.3.3 to detect points B and C. The features of each contour point are the x, y coordinates and the curvature at this point. The coordinates are translated to the coordinate system in Fig.2.17(c) with the origin at the lower left corner of the bounding rectangle, and are normalized by the length and width of the rectangle respectively. The membership functions for the points B and C are obtained through training of a small sample. An overall score which is the weighted summation of each membership function is used to select the landmark point from all the points. The weight of the curvature is set higher than the other two features because B and C have distinct curvature features.

Point A is more difficult to detect because the curvature of A is not as distinct as that of B and C, so we detect point A based on the positions of B and C. Since point A is always to the left of B, we only search within a subset of the points whose x coordinates are smaller than that of point B. Besides the three features used for points B and C, an additional feature is defined for point A, which is the curve length between A and C. An overall score of four membership functions is used to select point A. Since the curvature of A is not particularly distinct, equal weight for each membership function can serve the purpose here. The final seed contour (Fig.2.16(c)) is constructed by connecting A and B. We record the positions of A and B before contour deformation. Any new points generated during the deformation which are in between the two positions are set inactive, so that the contour will not be attracted to the fornix again.

In case there is no fornix connected to the CC, point C can hardly be found. This means the overall score obtained from the membership functions of point C is very low for every contour point. We set a minimum threshold (0.2 in our experiment) for the score. In the detection of C, if all scores are less than the threshold we believe there is no fornix.

For the thalamus, a point inside the thalamus is recognized indirectly from the ventricle. We trace down a certain distance from the center of each ventricle to locate a seed point inside each thalamus. The distance can be equal to the length or the height of the ventricle, and either way can usually make the point inside the thalamus. We perform k-means again with eight clusters, and the cluster that contains the seed point is the thalamus cluster. Because the intensity of the thalamus is not distinct from the intensities of its neighboring structures, the resulting cluster may leak out of the thalamus boundary. To overcome this, we

shrink the cluster using a sequence of morphology iterations interlaced with connected components. The root node for the connected components algorithm at each iteration is the seed point. Thus, during erosion, the cluster converges to the seed point while removing parts of the cluster that are not part of the thalamus. In our implementation, the number of morphology iterations is based on the size of the image. To complete the seed, we dilate the cluster the same number of steps to grow the seed back to a larger percentage of its original size (Fig.2.16(g)).

2.3.5 Results

We perform our algorithms on different sagittal MR images for CC segmentation and axial images for thalamus segmentation. Fig. 2.18 shows some final results. The image intensity varies a lot among these images, but it does not affect the results since our algorithm is independent of the image brightness.

To quantitatively evaluate the accuracy of our method, we compare our results (C_{seg}) with the ground truth (C_{true}) which is the manual segmentation by a trained expert. The measurements in equations (2.9)~(2.13) are used to calculate the similarity between our results and the manual segmentation. The experiment is performed on 20 CC images and 10 thalamus images (20 thalami), and the mean and standard deviation of each measurement on the CC and the thalamus are listed in Table 2.5.

Especially, the results on the CC segmentation are comparable with the results in 2.1, but this method does not require careful user initialization as in the semiautomatic method in 2.1. Overall, this algorithm has a better performance on the CC than on the thalamus because the CC boundary is more distinguishable, but both results can show high accuracy of this algorithm.

Table 2.5. Quantitative validation results

	CC		Thalamus	
	Mean	Std	Mean	Std
FNF	0.08	0.06	0.10	0.09
FPF	0.01	0.09	0.05	0.11
Dice	0.92	0.05	0.90	0.07
TPF	0.90	0.06	0.89	0.09
Overlap	0.84	0.08	0.77	0.10
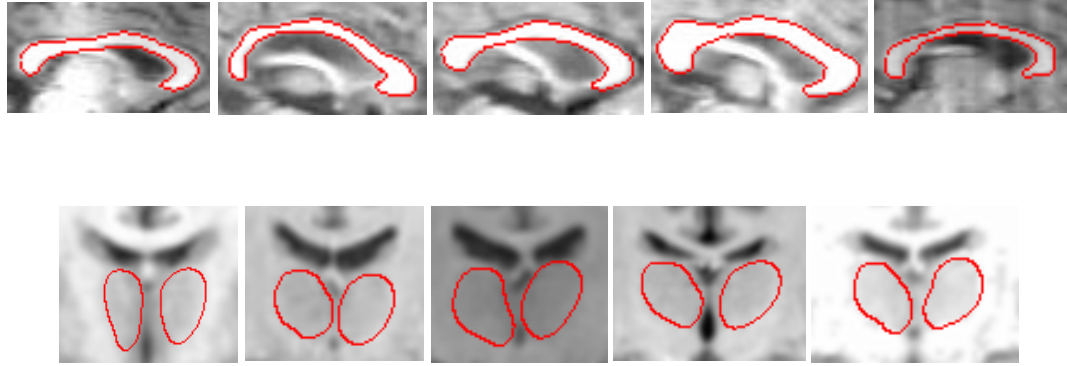
Fig. 2.18. Final results on different images (top: CC; bottom: thalamus)

2.3.6 Conclusion

This section describes an automatic segmentation framework which applies a hierarchical inference scheme to the initialization of the deformable models. In this way we can produce a seed that is more adapted to our target shape, which can greatly improve the speed and robustness of deformable models. The application of this framework to the CC and the thalamus is illustrated. The details of the segmentation of the two structures are quite different, but they follow the same framework. The quantitative validation shows the accuracy of our algorithm.

2.4 Summary

Three different methods are proposed in this chapter for brain structure segmentation. The first method (section 2.1) is targeted to the corpus callosum segmentation, the second method (section 2.2) can be used for any generic brain structure, and the third method (section 2.3) is a general framework with the application on two specific brain structures. All these methods are based on deformable models, and the main difference among them lies in the model initialization. Since classical deformable models often suffer local minima problem if the initialization is poor, the main contribution of this chapter is developing effective initialization schemes which bring the model close to the target at the initial stage, so that the subsequent deformable model can proceed as desired.

Implicit representation of deformable models is generally preferred to explicit representation because of its convenience in handling topological changes, especially in 3D case (section 2.2). However, explicit models are still applicable in some simple 2D topology with appropriate initialization (section 2.1 and 2.3). The first two methods are semiautomatic which require user intervention at the initialization stage, while the

third method is fully automatic. Although the fully automatic method saves the user operation, the semiautomatic method is sometimes preferred in clinical applications because it gives the doctors some degree of control over the segmentation. Therefore, there is no favor on either type of models or degree of automation, and the choice of methods depends on the specific application. Chapter 4 will show the different applications of our segmentation methods in the shape analysis of brain structures.

3. Point Correspondence in 3D Shape Analysis

In this chapter, we propose two methods for establishing 3D shape correspondence. This is an essential step in statistical shape analysis, and the case for 3D shapes is much more complex than for 2D shapes. Therefore, we focus on 3D shapes and the methods for 2D shapes can be easily derived. Section 3.1 describes a mapping based method which extends the classical conformal mapping technique to the mapping on a planar domain. Section 3.2 describes a non-mapping based method which employs landmark sampling and surface normal constraints.

3.1 Planar Conformal Mapping

In this section we propose a mapping based method. As stated in 1.3, spherical mapping cannot handle non-sphere-like shapes. For example, sometimes it is more convenient to represent a brain structure as an open surface, such as the corpus callosum (CC) surface in Fig 3.1 which is stitched from many 2D contours. In this case, the spherical parameterization is not practical because of extreme distortion. Conformal parameterization has been explored intensively as a potential approach to the matching and analysis of brain data. It has a flexible parameterization in that it can map the surface to different types of simple geometry. We design a planar conformal mapping method for the surface with two openings at the ends. The CC is used as an example for illustration, but this method can be applied to any open surface that is reconstructed from 2D contours.

In order to find point correspondence among different CC surfaces, we first parameterize each CC surface using a planar domain, and then align the parameterization in the planar domain so that surface points across different objects posses the same parameterization. Finally, the aligned parameterization in the planar domain is mapped back onto the original surfaces (Fig. 3.1).
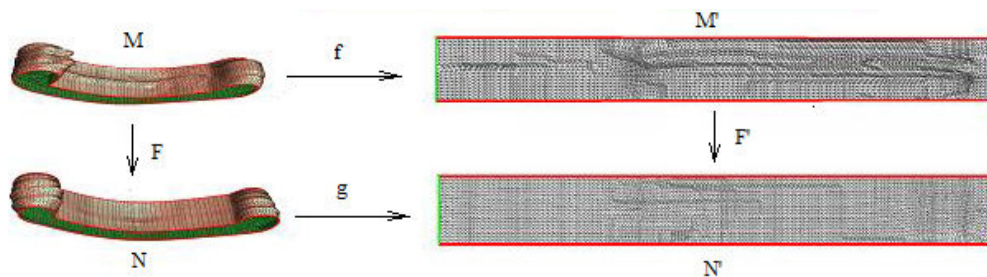
Fig. 3.1. Illustration of establishing point correspondence via planar conformal mapping: two individual CC models (left) and their planar domains (right)

3.1.1 Conformal Parameterization

The algorithms for global conformal parameterizations are directly inspired by Riemann surface theory [Jost 2000] and Hodge theory [Schoen and Yau, 1997]. Essentially, finding a global conformal parameterization for a point-set surface S can be intuitively interpreted as the computation of a holomorphic one-form. Given a scalar function $f : M \rightarrow R$ defined on manifold M, the differential of f, denoted as df, is a harmonic one-form if f is harmonic, i.e., $\Delta f = 0$. The complex form $\omega_1 + \omega_2 \sqrt{-1}$ is a holomorphic one-form if ω_1 and ω_2 are harmonic one-forms. After we obtain a holomorphic one-form $\omega_1 + \omega_2 \sqrt{-1}$, we map the surface S to the parametric plane by integration. Suppose that (θ_1, θ_2) are their parametric values. First we fix a base vertex v_0, for any vertex v_k, we select a curve γ on S from v_0 to v_k, and define the parameter of v_k to be

$$(\theta_1(v_k), \theta_2(v_k)) = \int_\gamma (\omega_1, \omega_2) \qquad (3.1)$$

Given a holomorphic 1-form (ω_1, ω_2), a horizontal trajectory is a curve on the surface S which is mapped to the horizontal curve (iso-θ_2) on the parameter plane by integrating the holomorphic one-form. A vertical trajectory is a curve on S which is mapped to the iso-θ_1 curve on the plane. The intersecting horizontal and vertical trajectories form the conformal net, which locally has a tensor product structure. More detailed discussion can be found in [Jin et al., 2004].

In order to obtain the conformal parameterization of the CC surface M on a planar domain M', a starting point needs to be selected on M. We choose the leftmost point on the lower boundary of the CC as the starting point. This point is a feature point on the CC surface which can be identified consistently on different CC surfaces, so it is a reasonable starting point to align all the CC surfaces. If we cut the surface M at the starting point along the direction perpendicular to the boundaries of M, the surface M is unfolded into a strip with the two boundaries γ_1 and γ_2. This strip is to be mapped onto the rectangle M' with γ_1 and γ_2 mapped to the upper and lower long edges of M'. A special holomorphic one-form $\tau + \eta \sqrt{-1}$ is

49

constructed such that its integration along the horizontal trajectory (iso-θ_2 curve) has a real part with a

constant period, and its integration along the vertical trajectory (iso-θ_1 curve) has a constant imaginary part.

To get the imaginary part η, we need to compute a harmonic function f by solving a Dirichlet problem:

$$\Delta f = 0,$$
$$f \mid_{\gamma_1} = 1, \qquad (3.2)$$
$$f \mid_{\gamma_2} = 0.$$

η is the gradient of f, and the real part τ is the conjugate of η. In our implementation, we first convert η

into a piece-wise constant vector field υ_η sampled at each face; rotate υ_η by $90°$ around the face normal,

and turn the rotated vector field into the one-form τ. With the starting point chosen before, we can

integrate the holomorphic one-form $\tau + \eta\sqrt{-1}$ on M. Taking the real and imaginary parts of the

integration to be y and x coordinates respectively, we get a horizontal strip which is periodic along the X

direction. Cutting the strip along the vertical line $x=0$, we get the final planar domain M', where the upper

and lower side corresponds to the two boundaries in the original mesh M. The map from M to M' is

conformal; the height-width ratio of the rectangle is a conformal invariant of the original surface. In our

experiments, we always rescale the rectangle uniformly along the X and Y directions so that the width

equals 2π. As a consequence, the intrinsic structure is solely reflected by the height of the rectangle for

our cases.

3.1.2 Aligning the Planar Domain

Suppose we have two different CC surfaces M and N. First, each surface is parameterized to its planar

domain M' and N' through the mapping function f and g, and one of them (e.g., M) is randomly selected as

the target surface. After that, we compute a mapping function F' between planar domains M' and N'. In our

case, a reasonable map should satisfy two constraints. First, each boundary of the source surface should be

aligned with the corresponding boundary of the target surface. In the rectangle domain, we need to map the

upper (lower) edge of the source rectangle to the upper (lower) side of the target rectangle. As mentioned

before, the width of the rectangle is always equal to 2π, while the height may be different from model to

model. Therefore we may need to rescale one rectangle along Y direction to align with the other rectangle.

In the case of the CC surface, all the rectangles have similar heights, so the rescaling factor is very close to one and therefore would not hurt the conformality of the map too much. Even in the case of a bigger rescaling factor, the map is still affine and therefore harmonic, which is already strong enough for our purpose.

The second constraint is to align the "bending corner" (i.e. the anterior most of the CC) across two surfaces. Since we choose the bending corner to be the starting point used in the integration of the holomorphic one-form, in the planar domain the bending corner will appear at the vertical cutting side. Aligning the two rectangles along their left and right sides, the bending corners will be automatically matched. Since all the rectangles have the same width 2π, we do not need to rescale any of them along X direction.

3.1.3 Aligning the Original Surfaces

Once the source rectangle N' is totally aligned with the target one M', we get a map F' between two planar domains. The map F' between two planar domains induces a map F between the two original surfaces M and N,

$$F = g^{-1}F'f \quad (3.3)$$

For each vertex v on the target rectangle, there is a point $F'(v)$ on the source rectangle, which may lie inside a triangle, on an edge or at a vertex. Without loss of generality, we represent the point as a linear combination $F'(v) = b_0 v_0 + b_1 v_1 + b_2 v_2$, where v_i is the vertex of the resident face on the source rectangle and b_i the corresponding barycentric coordinate. The map F' between two planar domains induces a map F between the two original surfaces. Namely, each vertex v in the target surface corresponds to a point $F(v) = b_0 v_0 + b_1 v_1 + b_2 v_2$ in the source surface. To facilitate later processing, we resample the source surface using these points as vertices, and re-triangulate the source surface with the same connectivity as that of the target surface. Therefore, each point on the source surface has a correspondence on the target surface, and the two surfaces have the same triangulation.

For a group of CC surfaces, one of them is selected as the target surface and the same method used to align each source planar domain to the target planar domain, and to re-triangulate the source surface. In this way, the point correspondence among all the surfaces is established through the target surface.

51

3.2 Landmark Sampling Based Method

A non-mapping method based on [Dalal 2007] is proposed in this section. The CC surface is again used as an example for illustration, but the CC surface is different from that in 3.1 in that it is a closed surface. This surface is directly obtained from 3D segmentation instead of stitching a set of 2D contours.

3.2.1 Spatial Alignment and Landmark Sampling

The processing in this section is very similar to those in [Dalal et al., 2007], and we only give a brief description here. The input is two triangle meshes, the template U and the target V. For the correspondence among a group of shapes, we can randomly choose one as the template, and correspond each of the other shapes to this template using the same method. The algorithm begins with a spatial alignment. Given two triangle meshes U and V, we first remove the location difference by moving their centers of mass to the origin. Then we normalize each shape by their centroid size [Bookstein 1989] so that there is no size effect in the following shape analysis. The rotation between the two shapes is removed by aligning their principle axes. The details of the rotation alignment can be found in [Dalal et al., 2007].

After alignment we construct a set of landmarks on the template shape U. These landmarks need not to be anatomically defined. The basic consideration is that they should be dense enough to represent the surface U and sufficiently sparse for a compact statistical shape model [Dalal et al., 2007]. The 3D space is divided into equal-sized cubic cells. In each cell, the surface point closest to the center of the cell is selected as the landmark [Dalal et al., 2007]. The size of the cell controls the density of landmarks. Typically, there are thousands of points on a 3D mesh, but the computational cost will be expensive if we include all points in the shape analysis. Only a few hundred of landmarks are sampled and used for shape analysis purpose. A comparison of the original surface points and the sampled landmarks can be seen from Fig.3.3(a) and (c).

3.2.2 Initial Correspondence on the Target Shape

The initial correspondences on the target V is constructed in a simple way in [Dalal et al., 2007]. Specifically, for each landmark on the template, a point on the target triangle mesh which is closest to this landmark is selected as the target landmark. Since the two shapes are pre-aligned, it is assumed that this simple method can find a rough correspondence between the landmarks on U and V [Dalal et al., 2007]. Generally, this simple method can work well for many types of shapes. However, if the two aligned shapes

do not have a certain amount of overlapping, this method may fail to find the correct correspondence. The following example of the corpus callosum shape will illustrate this problem. Fig.3.2 shows two corpus callosum shapes after spatial alignment. For simplicity, only 2D shapes are shown and the 3D case is similar. The major parts of the bodies of the two shapes do not overlap. For point A on the template (Fig.3.2), its closest match on the target shape is point A ', but obviously this correspondence is not correct because A is on the lower body of the corpus callosum and A ' is on the upper body. In this case, it will be very difficult to correct it in the following landmark refining. This problem is very common since there is large variation among corpus callosum shapes.

In order to solve this problem, we propose an improved initial correspondence algorithm with a mutual selection strategy and normal constraints. We observe that if the two points are corresponded points on two shapes, their surface normal directions should not differ too much. If their surface normal directions are nearly opposite (e.g, A and A ' in Fig.3.2), they are not likely to be the corresponded points. We use the constraint of normal directions in addition to the distance in the landmark selection procedure. Let the set of template landmarks obtained in 3.2.1 be U_L each template landmark $u_i \in U_L$, the original point set on the target mesh be V , each target point $v_i \in V$, and the set of target landmarks be V_L . Our algorithm can be outlined as follows.

1. For each $v_j \in V$, find a landmark $u_i \in U_L$ which is closest to v_j . Usually many target points will find the same template landmark u_i because the points in V are much denser than in U_L .

2. For each $u_i \in U_L$, record all the v_j 's that select u_i as the closest point. Denote this set of v_j 's as V_i . The points in V_i are the candidate points to match the landmark u_i .

3. If u_i has more than one point in its V_i , find a point v_j from V_i whose normal direction is closest to the normal of u_i . Remove v_j from V and add it to V_L . If u_i has no points in its V_i , find a point v_j from V that has the closest distance to u_i . Remove v_j from V and add it to V_L .

The resulting points in V_L are the corresponding landmarks on the target shape. The index of each point in V_L is the same as the index of its matched point in U_L , so the correspondence between U_L and V_L is

53

established. Fig.3.2 again illustrates how our algorithm overcomes the aforementioned problem. Target

points A ' and B both select A as the closest point in step 1, but A selects B as its correspondence

because the normal directions of B and A are close. No target point selects C in step 1, so C selects a

closest point as its correspondence which is A ' in step 3. Fig.3.3 shows a real example of the initial

landmarks on the target corpus callosum shape obtained using the algorithm in [Dalal et al., 2007] and our

algorithm respectively. Sagittal view of the 3D model is displayed. Two aligned original shapes are shown

in Fig.3.3(a) which is similar to the case in Fig.3.2. Only point clouds are shown here without triangle

meshes. The upper body of the target shape is missing in the result of the algorithm in [Dalal et al., 2007]

(Fig.3.3(b)). The landmarks in our results, however, can correctly represent the target shape (Fig.3.3(c)).
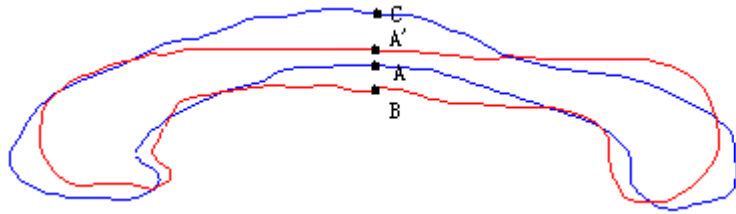
Fig.3.2 Illustration of the case where the algorithm in [Dalal et al., 2007] will fail. (blue: template; red: target)
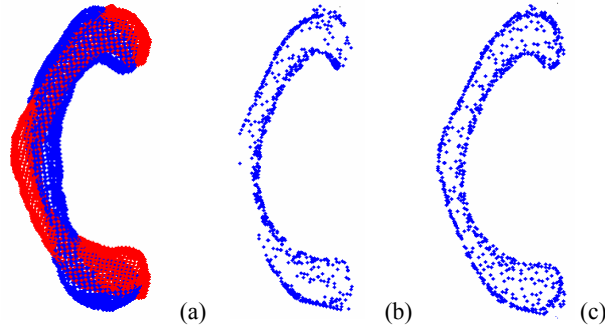
Fig. 3.3 (a) two aligned original point clouds overlaid (blue: template, red: target) (b) initial landmarks on the target

shape obtained using the algorithm in [Dalal et al., 2007] and (c) our algorithm

3.2.3 Refining Landmarks by Iterative TPS

Given the initial correspondence between two sets of landmarks, U_L and V_L , one can estimate a

transformation T from V_L to U_L , and T can be applied to the entire target point set V . In [Belongie

2002], TPS transformation is used iteratively to refine the shape context matching. Similarly, we can

perform our algorithm in 3.2.2 and TPS transformation iteratively to refine the shape correspondence. We

will now briefly introduce the basic concept of TPS. Suppose there are n landmarks in U_L and V_L respectively. Let the landmarks in U_L be $u_i = (u_{ix}, u_{iy}, u_{iz})$, and the landmarks in V_L be $v_i = (v_{ix}, v_{iy}, v_{iz})$, $i = 1...n$. u_i and v_i are corresponded points for the same i . The TPS function which transforms v_i to u_i has the form:

$$f(v_{ix}, v_{iy}, v_{iz}) = a_0 + a_x v_{ix} + a_y v_{iy} + a_z v_{iz} + \sum_{j=1}^{n} w_j \phi(\| (v_{jx}, v_{jy}, v_{jz}) - (v_{ix}, v_{iy}, v_{iz}) \|) \quad (3.4)$$

where $\phi(r) = r$ in 3D case. We use three separate functions to model the transformation of x,y,z coordinates, and the objective function is

$$T(v_{ix}, v_{iy}, v_{iz}) = (f_x(v_{ix}, v_{iy}, v_{iz}), f_y(v_{ix}, v_{iy}, v_{iz}), f_z(v_{ix}, v_{iy}, v_{iz})) = (u_{ix}, u_{iy}, u_{iz}) \quad (3.5)$$

Each function f_x, f_y, f_z has the same form as (3.4). Equation (3.5) can be written in the matrix form and the coefficients a_0, a_x, a_y, a_z, w_j of each function can be solved from the linear equation.

After we obtain the initial correspondence between U_L and V_L , we estimate a TPS transformation from V_L to U_L , and apply the transformation function to the entire point set V . The transformed target point set is denoted as V^1 , and the algorithm in 3.2.2 is then applied on U_L and V^1 to find a set of target landmarks V_L^1 . Then a TPS transformation from V_L^1 to U_L is estimated and applied to V^1 . The final landmark set V_L^n after n iterations contains the landmarks in the transformed shape V^n . We keep track of the indices of the landmarks in each iteration, so that the final landmarks can be easily mapped to their original coordinates in V . After several iterations, the initial correspondence error can be greatly reduced. As suggested in [Belongie 2002], three iterations are used in our experiment. Fig.3.4 shows one example of landmarks before and after three TPS iterations. We can find that the surface points are more evenly distributed after TPS iterations.

While the landmark sliding in [Dadal et al., 2007] can obtain similar results of refined landmarks, our method requires less computational time, and all the sampled landmarks are on the original surface without sliding away.

Fig.3.4 The landmarks on a target shape before (left) and after (right) TPS iterations

3.3 Summary

The 3D shape correspondence problem is a broad topic and cannot be all covered in this thesis. The main contribution of this chapter is the methods to solve the 3D shape correspondence problems in specific cases. The first method (section 3.1) is particularly designed for open surfaces with two boundaries. Since it is a mapping based method, the parameterization can be performed on the original surface points thus maintaining the high resolution of the original surface. The second method (section 3.2) aims at solving the mismatching problem due to individual shape variations among different shapes. It is conceptually simple but can effectively solve the mismatching problem. It can be applied to both open and closed surfaces although we only show the example of closed surfaces. The resolution of the surface points is lower after the correspondence is established because down-sampling is involved, i.e., the correspondence is only built on a subset of the original surface points.

As stated in 1.3, the first method is usually used for whole surface analysis because it is faithful to the point density of the original surface, while the second method is not always a good choice for whole surface analysis due to the loss of point resolution. However, the second method can be used for pseudo-landmark analysis with additional constraints of anatomical landmarks, which is still preferred in most clinical applications as long as the requirement of surface resolution is not critical. Chapter 4 will show more clinical applications of the two methods.

4. Quantitative Analysis of 3D Brain Structures in Clinical Applications

This chapter gives several examples of clinical applications. The methods presented in Chapter 2 and 3 are applied to perform quantitative analysis of certain brain structures. Section 4.1 performs landmark analysis and pseudo-landmark analysis of the 2D corpus callosum in autism. Section 4.2 performs global traits analysis and whole surface analysis of the 3D corpus callosum in autism, where the corpus callosum model is an open surface reconstructed from 2D contours. Section 4.3 performs similar analysis as in 4.2 for the case of Phenylketonuria while the corpus callosum model is a closed surface. Section 4.4 performs pseudo-landmark analysis of 3D ventricles in order to find the shape difference between genders.

4.1 Quantitative Analysis of the 2D Corpus Callosum in Autism

4.1.1 Background

Autism is a severe developmental disorder characterized by social deficits, impaired communication, and restricted and repetitive behavior patterns [American Psychiatric Association, 2000]. Brain MRI study has reported inconsistent results regarding the abnormalities of brain structures in autism. The inconsistency may be due to factors such as the sample size, subject age and gender. Heterogeneity within the autism diagnosis can also significantly obscure the genetic basis of the disorder [Miles et al., 2005].

The corpus calllosum (CC) is the major commissural pathway between the brain hemispheres and plays an integral role in relaying sensory, motor, and cognitive information from homologous regions in the two hemispheres. Many studies have reported abnormalities of the CC in autism. Quantitative morphologic assessment of individual brain structures is often based on volumetric measurements and shape analysis. Volume comparison gives global information of the size difference between pathological and healthy structures, but no local shape difference is revealed. Early works focused mostly on the area of the CC on the midsagittal slice [e.g., Piven et al., 1997; Hardan et al., 2000], and most of them found reductions in the size of different sub-regions of the CC in autism. Shape analysis, on the other hand, can precisely locate morphologies at any location on the brain structure. However, point correspondence among individual shapes is a crucial and difficult problem. [Vidal et al., 2006] found the anterior and posteriors of the CC were more inward in autism. Point correspondence was established globally across the entire shape. Note

that their shape correspondence was based on 2D contours of the CC on sagittal MR images, because the shape characters of the CC can be sufficiently reflected by the sagittal view.

In this section, we conduct both landmark analysis and pseudo-landmark analysis to examine the abnormalities of the CC in autism. A configuration of landmarks was identified in brain MRI midsagittal sections based on a predefined criterion [Ozdemir et al., 2007]. EDMA and TPS procedures are performed in the landmark analysis. Pseudo-landmarks are generated in between landmarks, and statistical methods were used to compare two groups of shapes (autism vs. control) at every location in the global shape comparison. In addition, a new local shape comparison is proposed which compared each part of the shape in its local coordinate system.

4.1.2. Participants

Twenty-five children with autism were recruited from the Thompson Center of Autism and Neurodevelopmental Disorders. All patients came from families who had come to Thompson Center and gone through our standard diagnostic protocol (ADI-R, ADOS, cognitive, genetics). Children with disorders known to cause an autism phenotype such as Fragile X, chromosome abnormalities or severe prematurity with brain damage, children with IQs/DQs less than 50 and children with premature puberty were excluded. To maximize homogeneity this essential autism group were further limited to individuals with no history of seizures, abnormal brain EEGs, or abnormal MRIs.

Eighteen control subjects were recruited from the community under the regulations of the Thompson Center control subject recruitment protocol. The control group of typically developing children were matched for gender, age and ancestry. They underwent a short intake history to rule out significant language, cognitive or social delays. Children receiving special education with an IEP, diagnosed and treated ADHD symptoms and other childhood psychiatric disorders and children with a sibling diagnosed with autism were excluded. Table 4.1 summarizes the demographic characteristics of the autistic patients and comparison controls. Student t-test was used to compare the ages and χ^2 test was used to compare the gender ratios. There is no significant difference between the two groups in terms of age, gender and race.

This study was approved by the Health Sciences Institutional Review Board. The parents or legal guardians of all subjects provided written consent for participation in this study, while the subject provided voluntary assent.

Table 4.1 Demographic characteristics of autistic patients and controls

Measures	Patients	Controls	Test statistics	P value
Age (mean±std)	6.86±2.87	8.44±4.34	t=0.861	0.197
	(range 3.6~12.8)	(range 4.5~14.2)		
Gender (M:F)	20:5	14:4	χ^2= 0.030	0.975
Race	All Caucasians	All Caucasians		

4.1.3 CC Segmentation and Landmark Collection

Axial, coronal and sagittal T1-weighted images were acquired using the Siemens Symphony 1.5 T scanner with the following parameters: TR = 35ms, NEX = 1, flip-angle = 30 degrees, thickness = 1.5mm, field of view = 22cm, matrix =512x512. Sedations were performed if needed on some of the patients based on our autism aesthesia protocol. 16-bit MR data were compressed to 8 bits by linearly rescaling the voxel intensities. Any voxel whose intensity was below the 2nd percentile was set to 0, and any voxel whose intensity was above the 98th percentile was set to 255. The remaining voxel intensities were then linearly interpolated between these two extremes to ensure that all data points lie within the 8-bit scale. The data were aligned with the stereotactical coordinate system and re-sliced to isotropic voxels of 1mm^3 using Slicer (www.slicer.org).

From the sagittal planes, the midsagittal section that most clearly displayed the cerebral aqueduct, CC, and superior colliculus, was selected manually [Ozdemir et al., 2007]. From the midsagittal image, the CC was segmented using the semiautomatic method in 2.1. The segmentation was performed by the same trained expert for all the data. After the CC contour of each subject was extracted, they were aligned in order to remove the shape differences due to translation, scaling and rotation. The aligning method is the same as in 3.2.1. The centroid size obtained from the scaling step is the only size measurement that is uncorrelated with shape variation [Bookstein, 1991]. Therefore, our analysis reveals pure shape difference without the effect of the size.

Nine anatomical landmarks were identified on each of the aligned CC shapes (Fig. 4.1). Anatomical landmarks are biologically meaningful loci that can be repeatedly located with high accuracy and precision [Richtsmeier et al., 1995]. We followed the landmark definition in [Ozdemir et al., 2007] which included extreme points or terminals and maxima of curvature (Table 4.2). All landmarks were manually identified

by a single rater. Intra-rater and inter-rater reliability of landmark selection was tested using the method in [Ercan et al., 2008]. Two raters each performed the landmark selection twice, and G-coefficients between two landmark sets from the same rater (2 pairs) and from two different raters (4 pairs) were calculated. A G-coefficient close to 1 indicates high intra/inter-rater reliability. The minimum G-coefficient for all six pairs of landmark sets was 0.994, which indicated high reliability of the landmark selection.
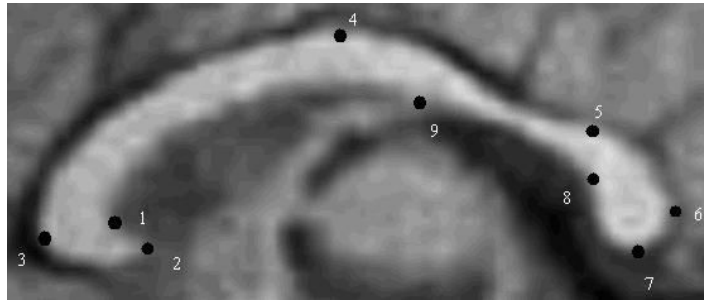
Fig.4.1 Landmarks of the CC

Table 4.2. Landmark definitions of the CC

Landmarks	Definitions
1	Interior angle of genu
2	Tip of genu
3	Anterior most of CC
4	Topmost of CC
5	Splenium topmost point
6	Posterior most of CC
7	Bottommost of splenium
8	Interior notch of splenium
9	CC-fornix junction

4.1.4 Landmark Analysis

In this section only anatomical landmarks are used for the analysis. EDMA and TPS analyses are performed.

In EDMA analysis the shape change was measured in terms of the overall landmark configuration. Euclidean distances were calculated between every pair of landmarks, leading to a form matrix for each shape [Burrows et al., 1995]. k landmarks had $k(k-1)/2$ inter-landmark distances. Since the form matrix is symmetric, a form vector of length $k(k-1)/2$ can be used to describe the matrix. Box's M-test [Dryden and

Mardia, 1998] was performed to test the homogeneity of the variance. EDMA I method was used if there was existing homogeneities of variance–covariance matrices, otherwise, EDMA II method was preferred for shape analysis.

We adopted EDMA I method because the hypothesis of homogeneity of variance could not be rejected (p=0.95). We performed statistical tests on the null hypothesis that the two groups of shapes did not differ in form [Burrows et al., 1995]. In brief, the average form vector was computed within each group, and the ratio of the two average form vectors was the form difference vector. The raw test statistic was computed from the maximum value (max) and the minimum value (min) of the form difference vector, which was $T_0=max/min$. Permutation test was used to find the statistical significance of the difference between the two form vectors. In this procedure, a bootstrap was used to generate M permutated samples and thus M test statistics (T^*) can be computed in the same fashion as T_0. The percentage of T^* which were greater than T_0 was the p-value. A significant level of 0.05 was selected and $M=1000$ in our experiment.

We also performed traditional t-tests on each inter-landmark distance under the null hypothesis that each distance was the same in the two groups. Since only a small number (9*8/2=36) of distances were tested simultaneously, Bonferroni procedure [Shaffer, 1995] was used to adjust each p-value for multiple comparison.

In TPS analysis, the shape change is described by the deformation from one group of shapes to the other. The theory of TPS is described in 3.2.3 in its 3D form. Here we restate it in its 2D form. A mapping function $f(x,y) = [f_x(x,y), f_y(x,y)]$ maps the location (x,y) to a new location (f_x, f_y). Each of the functions f_x and f_y has the form:

$$f_*(x,y) = a_0 + a_x x + a_y y + \sum_{i=1}^{n} w_i \phi(\| (x_i, y_i) - (x,y) \|) \qquad (4.1)$$

where $\phi(r) = r^2 \log r^2$, (x_i, y_i) are the landmark coordinates on the starting form, * can be x or y, and n is the number of landmarks. Given that $f(x_i, y_i) = \vec{v}_i$ where $\vec{v}_i$ is the corresponding landmark on the target form, the coefficients a_0, a_x, a_y, w_j can be solved from a linear equation.

Define the following matrices

$$K = \begin{bmatrix} 0, \phi(r_{12}),..., \phi(r_{1n}) \\ \phi(r_{21}),0,..., \phi(r_{2n}) \\ .... \\ \phi(r_{n1}),......,0 \end{bmatrix}_{n \times n} \quad P = \begin{bmatrix} 1, x_1 y_1 \\ 1, x_2 y_2 \\ .... \\ 1, x_n y_n \end{bmatrix}_{n \times 3} \quad L = \begin{bmatrix} K, P \\ P^T, O \end{bmatrix}_{(n+3) \times (n+3)} \quad Y = \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \\ ... \\ \vec{v}_n \\ O \end{bmatrix}_{(n+3) \times 2}$$

where O is a matrix of zeros whose size depends on the parent matrix it is in and $r_{ij} = \| (x_i, y_i) - (x_j, y_j) \|$.

The two columns of the matrix $L^{-1} Y$ are the coefficients in the functions f_x and f_y respectively. The $2n$ coefficients w_j (n coefficients in each direction) correspond to the non-uniform (non-affine) transformation. The expansion factor at each landmark can be computed from the Jacobian of w_j with respect to the transformed landmarks [Bookstein, 1991].

Two types of shape deformation were calculated. The first type was the deformation from the mean landmark form of the controls to that of the patients. The second type was the deformation from the overall mean landmark form (both patients and controls) to the mean forms of the patients and the controls respectively. For each type we used the data analysis package PAST [Hammer et al., 2001] to calculate the expansion factors and display the deformation grid.

The overall mean landmark form was also deformed to the landmark set of each individual subject, and multivariate statistic analysis was used to compare the non-uniform coefficients w_j between patients and controls. Specifically, Hotelling T^2 two-sample test was performed separately on the coefficient vectors in x direction (n-D), the coefficient vectors in y direction (n-D) and the concatenated coefficient vectors of both directions (2n-D). A positive, larger T^2 corresponds to a smaller p-value, which indicates more significant group difference in the landmark deformation.

4.1.5 Pseudo-landmark Analysis

In this section, shape difference was examined at every location on the shape boundary. Two types of shape comparison were conducted: one in the global coordinates and one in the local coordinates. Point correspondence among individual shapes must be established before any shape comparison. Since the landmarks in Fig.4.1 represent certain anatomical features on the CC shape, we assume that they are already corresponded. The landmarks divide the contour into several segments, and pseudo-landmarks are

interpolated on each segment using a uniform sampling. The sampling can be implemented by axis parameterization in which the sampled points are equally spaced along the axis defined by the two terminal landmarks, or by arc-length parameterization in which the points are equally spaced along the contour. These two methods have similar results when the contour is close to a straight line. For axis parameterization, we require the contour segment to be a single valued function of the points on the axis in order to obtain a unique point on the contour at each sampling point on the axis. The contour segments defined by our landmarks satisfy this requirement, and there is no global curvature extreme within each segment because the landmarks already cover the curvature extremes on the contour. Therefore, we used axis parameterization since it was simple. Fig.4.2(a) illustrates the uniform sampling on one segment. For each segment, we sampled the same number of points on different shapes, so that these points were corresponded by their indices. For simplicity, we ignored landmark 9 in the sampling process because the segment between landmarks 1 and 8 already satisfied the single value condition. In order to keep the point density the same along the entire shape contour, we set the number of sampled points on each segment proportional to the average length of this segment across different shapes. Fig.4.2(b) shows the resulting point correspondence between two CC contours. The point correspondence can be used for both global comparison and local comparison.

In the global comparison, aligned CC shapes were compared point by point across the whole contour. Since each point coordinate is a 2D vector, Hotelling T^2 two-sample metric was used again. Each raw p-value obtained from the test statistic T^2 was an optimistic estimation because the comparisons were made at hundreds of CC contour points. It was important to control the p-values for the multiple comparison problem. Non-parametric permutation tests [Pantazis et al., 2004] and False Discovery Rate estimate (FDR) [Hochberg et al., 1995] were typically used for p-value correction. We adopted FDR because it provides an interpretable and adaptive criterion with higher power, and also computationally efficient [Styner et al., 2006]. FDR method allows the false positive to be within a small proportion α ($\alpha = 0.01$ in our experiment). The p-value correction is computed as follows.

1. Sort the p-values such that $p_1 < p_2 < ... < p_N$, where N is the number of vertices on the surface.

2. From all the p-values that satisfy $p_i <= \alpha \bullet i/N$, select the p-value with the largest index i, denoted as p_α.

3. Declare all locations with p-values less than p_α significant.

63

In the local comparison, we focused on the group difference in each part of the CC shape instead of the whole shape. Each contour segment was compared in its local coordinate system as shown in Fig.4.2(a). The x-axis was the line connecting the two terminal landmarks, and we scaled the distance between the two terminal landmarks to 1. In this way the same segments across difference shapes were aligned by a common local coordinate system. Since our uniform sampling for point correspondence was based on x-axis, the corresponded points had the same x coordinates in their local coordinate system. For each segment, we compared the y-coordinates of the points between the two groups using t-tests followed by FDR correction. While the global comparison gives the group difference of the whole shape regardless of the segments, the local comparison gives the shape difference with regard to each part of the shape.
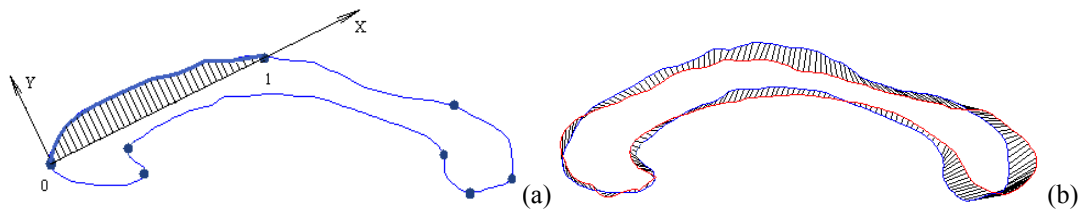
Fig.4.2 (a) uniform sampling and local coordinate system on one segment of the CC contour (b) the correspondence between two CC contours

4.1.6. Results

In EDMA analysis, no significant difference was found in the landmark form between the two groups of shapes (p=0.41). In the tests of individual distances, most of the inter-landmark distances did not show significant differences between two groups after p-value correction. The only significant difference was found in the distance from landmarks 1 to 6 (p=0.0006). Fig.4.3 shows the mean landmark configurations of the patients and the controls overlaid on the mean shape of the CC. The distance 1-6 was longer in the control shape than in the patient shape.

Fig.4.4 shows the TPS transformation grid along with the expansion factors of the transformation to each group from the overall mean. In the deformation from overall mean to the patients (Fig.4.4(a)), landmark 3 (anterior most) exhibits expansion and all other landmarks exhibit shrinking. Landmark 7 (posterior bottom) has the largest shrinking. In the deformation from overall mean to the controls (Fig.4.4(b)),landmark 3 (anterior most) exhibits shrinking and all others exhibit expansion. Landmark 7 has the largest expansion. The deformation from the controls to the patients (Fig.4.4(c)) is similar to the

deformation in Fig.4.4(a) except the expansion and shrinking are stronger. Landmark 7 again has the largest shrinking.

Hotelling T^2 test indicated that there was significant difference between patients and controls in the concatenated deformation coefficients (p=0.01), but no significant difference in the coefficients of x (p=0.09) or y direction (p=0.9) alone.

Fig.4.5(a) shows the average CC shapes of the two groups, which gives a descriptive visualization of the shape difference between two groups. The CC of the patients is more inward at both ends, resulting in a shorter distance in anterior-posterior length which is consistent with the results from EDMA. Note that the difference in the length does not reflect the size because we have removed the size difference in our spatial alignment. It is rather due to the different bending degrees of the CC body. Fig.4.5(b) and (c) show the raw and corrected significance maps of the global shape comparison (significance level=0.05). The raw p-values suggest significant shape difference in the anterior most and top, anterior lower body and posterior, while the corrected p-values retain significant difference in the anterior lower body and posterior bottom. This is consistent with the results from TPS where the posterior bottom of the patients shape shows severe shrinking from the controls shape. Fig.4.5(d) and (e) show the raw and corrected significance maps of the local shape comparison, where landmarks are highlighted in a different color. There is significant shape difference in the anterior bottom and posterior lower body in the raw p-values, and the corrected p-values retain the significance in the anterior bottom and part of the posterior lower body (isthmus). Both anterior bottom and isthmus bottom show different bending degrees between patients and controls, which causes the local shape difference.
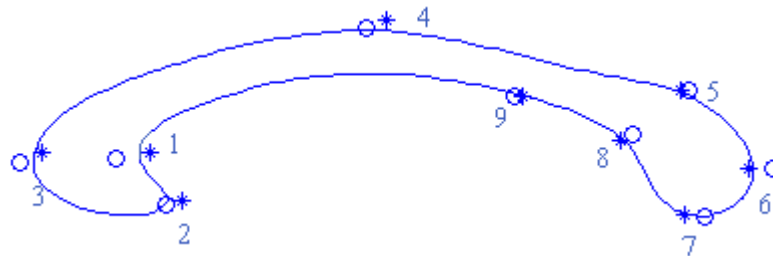
Fig.4.3 Overlaid mean landmark configurations of the patients (*) and the controls (o) on the overall mean shape.
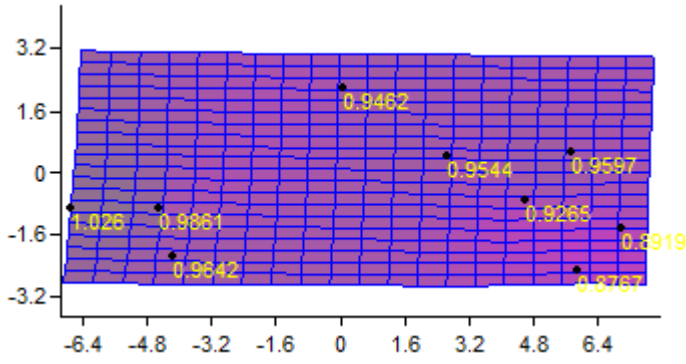
(a) overall mean to patients
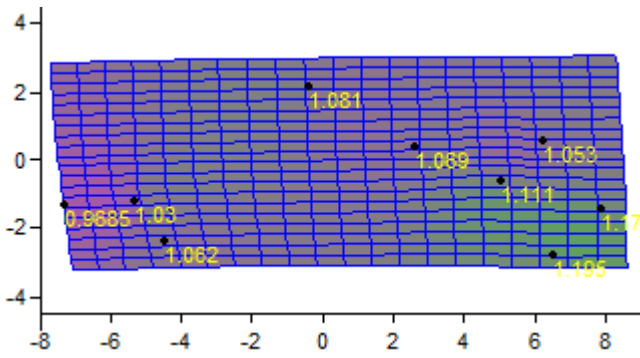
(b) overall mean to controls
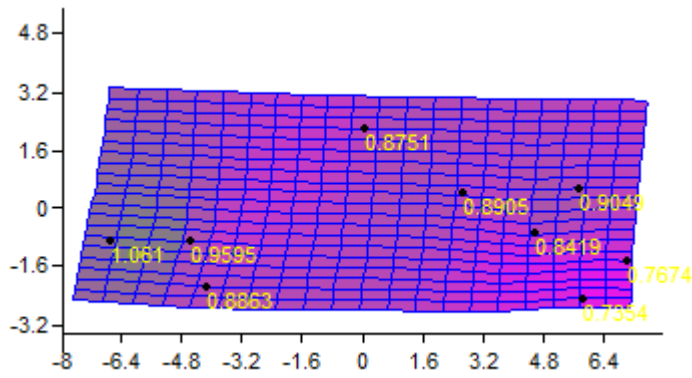
(c) controls to patients

Fig.4.4 TPS transformation grids and the expansion factor at each landmark

(a)

p=0   p>=0.05
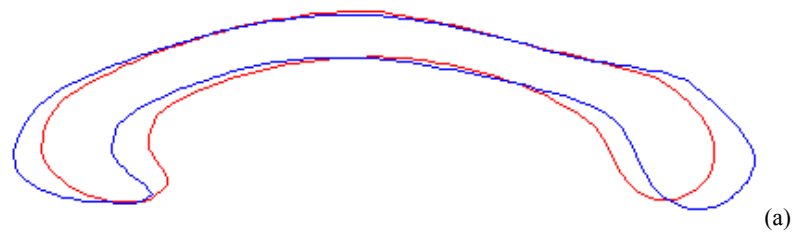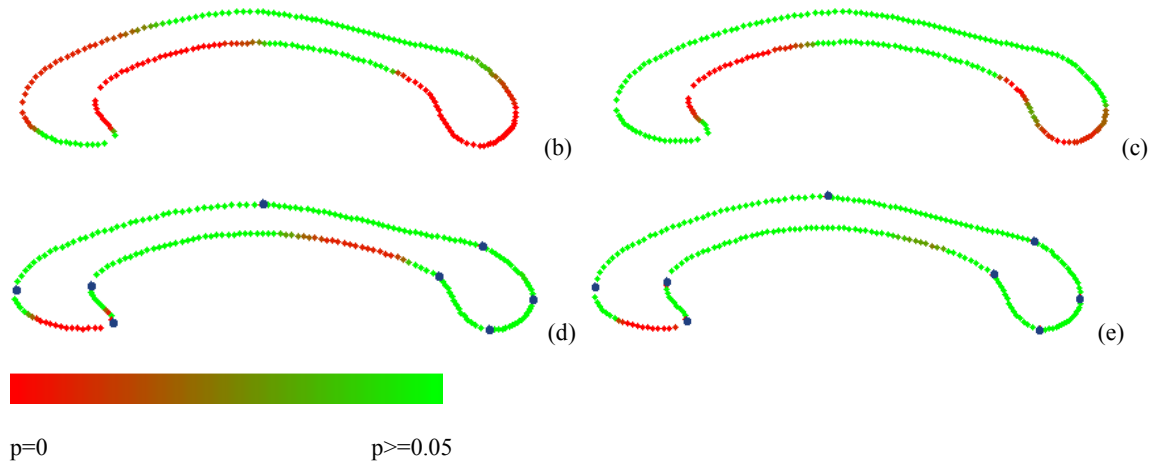
Fig.4.5 (a) average shapes of the CC (red: patients; blue; controls) (b) raw significance map of the global comparison (c) corrected significance map of the global comparison (d) raw significance map of the local comparison (e) corrected significance map of the local comparison (landmarks are shown as blue dots in (d) and (e))

4.1.7. Discussion and Conclusion

To our knowledge, this study was the first that used landmark based methods to analyze the CC abnormalities in autism, and the local shape comparison had not been performed before.

Our analysis permits some insight into the callosal functions potentially involved in the pathology of autism. The connections across the corpus callosum are topographically organized where the relay of sensory, motor and cognitive information is transmitted from two cerebral hemispheres. Existing findings have suggested the complexity of the callosal connectivity. For example, [Moses et al., 2000] examined the regional size reduction of the CC in children with focal lesions, and confirmed the cortico-callosal topography documented in adult and nonhumans. It also suggested limits to developmental neuroplasticity subsequent to perinatal brain injury. [Aboitiz et al., 1992] reported that primary and secondary sensory information was transmitted via large diameter callosal fibers and higher-order sensory and cognitive information was transmitted through small diameter fibers. There are different representations of large and small diameter fibers in different calllosal channels [Aboitiz et al., 1992]. Anterior callosal regions may be involved in the transmission of cognitive information [Clarke et al., 1998]. The isthmus is where the callosal motion fibers cross through [Wahl et al., 2007], and it has been demonstrated to connect myelinated fibers from posterior language regions and auditory association areas [Clarke and Zaidel, 1994]. Local shape differences are found in these areas in our study, which may be associated with the aberrant

67

cognition and impaired verbal communication in autism. Posterior callosal regions are involved in transmitting sensory information [Clarke et al., 1998], and autistic patients often have extreme sensory issues (hypersensitivity or hyposensitivity) which may be related to the more inward posterior of the patients found in our study.

The results of this study yet need to be interpreted with caution. We include both males and females in our subjects, which may overlook the gender difference of the CC [Ozdemir et al., 2007]. However, due to the unbalanced male-to-female ratio (4:1) in autism, it's hard to conduct separate experiment in female patients because of the sample size. We therefore use a matching ratio of male to female in the control subjects. The sample size in this study is relatively small, which may affect the statistical results.

The contribution of this study is two-fold. First, the findings of this study provide some insight in the pathology of autism related to the functions of the CC, although further studies need to be performed to confirm the results. Secondly, the procedures of our analysis can be applied to similar studies of other brain structures. The CC is a special brain structure whose shape can be characterized by a 2D contour, thus the methods used in this section are relatively simple and easy to be generalized to other 2D biological shapes. However, most brain structures are naturally 3D, which adds more difficulty to the shape analysis. The following sections will discuss the methods in 3D shape analysis and its clinical applications.

4.2 Detecting Corpus Callosum Abnormalities in Autism using Planar Conformal Mapping

4.2.1 Overview

This section gives a comprehensive pipeline of the analysis of the CC in autism with respect to the shape morphology and global traits (Fig.4.6). Since the CC model is constructed from stacked 2D contours, planar conformal mapping is applied to establish the point correspondence in the shape analysis. The entire procedure can be summarized in the following five steps.

1. For each subject, CC contours are segmented from mid-sagittal slice and 4 adjacent slices on each side, and a 3D model of the CC is constructed by contour stitching.

2. CC volumes and measurements of the oriented bounding box are computed, and differences of these measurements are compared between patients and controls using t-tests.

3. Each CC model is mapped onto a planar domain and re-triangulated to establish the point correspondence among different models.

4. Spatial alignment of all CC models is performed using Procrustes analysis.

5. 3D coordinates at each surface point are compared between patients and controls using Hotelling T2 tests, which results in a significance map of group differences.
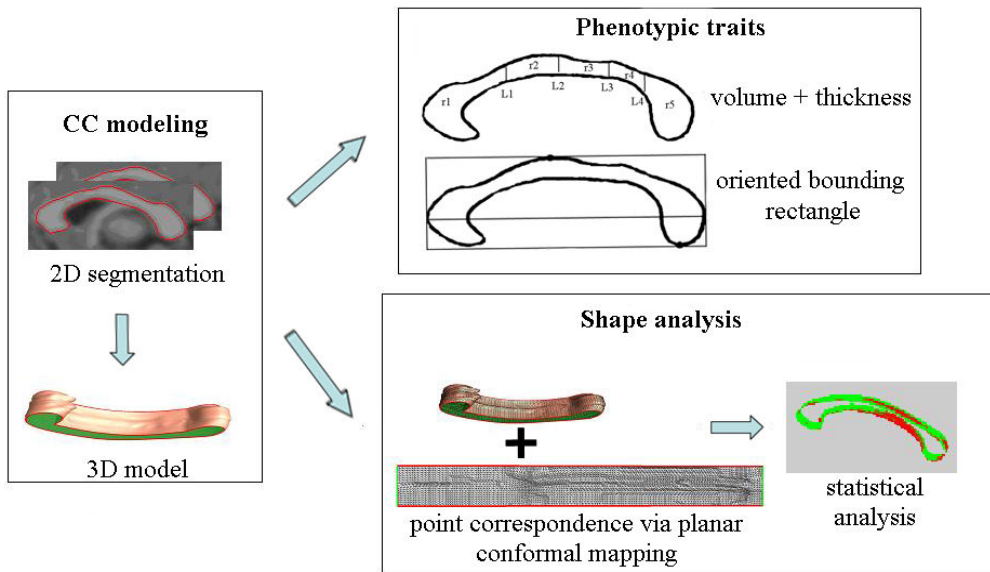
Fig.4.6 Schematic view of the entire procedure

4.2.2 CC Modeling

We start with slice-by-slice segmentation and stack up the 2D contours to create the 3D model of CC. Mid-sagittal slice and four adjacent slices on each side are used for CC segmentation. We use the semiautomatic method in 2.1 to segment the CC. Since the CC shapes on adjacent slices do not differ too much, we apply the segmentation method on mid-sagittal slice and the result is used as the seed for its two neighbor slices. After slight deformation the boundary curves on these two slices can be obtained and each one is again used as the seed of its next adjacent slice.

Contour stitching is performed to create the 3D CC model of each subject. A tangential Laplacian smoothing [Wood et al., 2000] is performed to maintain a good node distribution of the model. The reconstructed model is shown in Fig.3.2. The inner side of the surface is rendered in green in this figure.

4.2.3 Global Traits Analysis

69

Volumes of the CC and its sub-regions are computed similar to that in [Vidal et al., 2006]. The CC is divided into five regions along anterior-posterior line (Fig.4.7(a)). To find the anterior-posterior line, we calculate the distance between every two points on the CC boundary, and find the pair of points with the longest distance (p1 and p2 in Fig.4.7 (a)). The line connecting these two points is the anterior-posterior line. The five regions are defined the same as in [Vidal 2006]. For simplicity, we label them from r1 to r5 as shown in Fig.4.7 (a). Areas are computed in pixels for each 2D segmented CC, and the areas of 9 slices are summed to generate the voxel count of the 3D model. The raw volume is the multiplication of the voxel count and the voxel size in mm^3. To take into account the effect of the brain volume, we normalize the raw CC volume by the total brain volume (TBV) and the intracranial volume (ICV) respectively. TBV includes grey matter and white matter and excludes ventricles [O'Brien et al., 2006], and ICV is the sum of white matter, grey matter, and inner and outer cerebrospinal fluid spaces [Wolf et al., 2003]. A choice can be made between using TBV or ICV as an adjustment factor [O'Brien et al., 2006], but our results show that the two choices do not make any difference. We used the FMRIB software library [Smith et al., 2004] for brain skull stripping, and then used a web-based program [Karsch et al., 2008] to calculate TBV and ICV. The brain volumes, raw and scaled volumes of the CC are compared between patients and controls using t-tests.

We explore additional global traits based on the oriented bounding rectangle of the CC. To construct the rectangle, we first find points p1 and p2 as mentioned above. On each side of line p1p2, we find a point on the CC boundary which has the maximum distance to line p1p2 (p3, p4 in Fig.4.7 (b)). The bounding box is a rectangle whose long edges pass p3 and p4 respectively and are parallel to line p1p2, and whose short edges pass p1 and p2 respectively (Fig.4.7 (b)). We measure the length, width, and aspect ratio (length/width) of the bounding rectangle on each 2D slice and average them across the 9 slices. The shape of the bounding box depends on both the size and the shape of the CC, so these traits can reflect both size and shape information of the CC. We also compute the thickness at the border of every two sub-regions, which is the length of each dividing line (L1~L4 in Fig.4.7 (a)). With a little abuse of notation, we denote the thickness at each dividing border as L1~L4. The thickness at each border is averaged across the 9 slices. The average measurements of the bounding box and the thickness are compared between patients and controls using t-tests. To account for the brain volume effect, we scale the above raw measurements by the

70

cubic root of TBV and ICV respectively. Since these measurements are in mm, this scaling makes the units consistent. These scaled measurements are also compared between patients and controls.

(a)   (b)
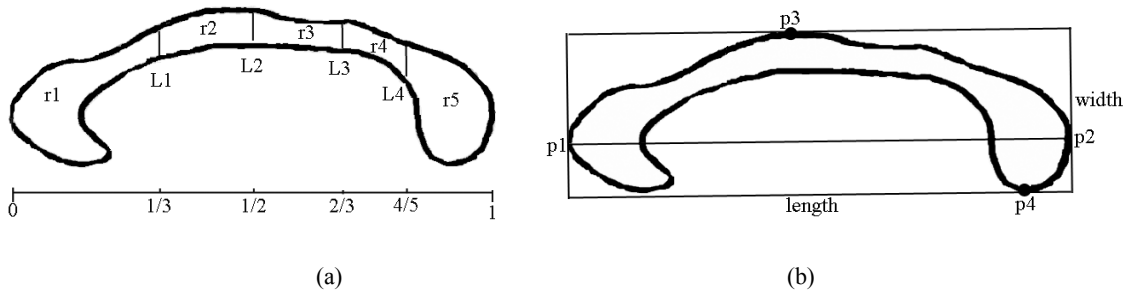
Fig.4.7. (a) partition of the CC: r1—anterior third; r2—anterior body; r3—posterior body; r4—isthmus; r5—splenium. (b) bounding rectangle of the CC

4.2.4 Statistical Shape Analysis

Planar conformal mapping described in 3.1 is used to establish the point correspondence among individual shapes. After that, generalized Procrustes Analysis [Gower 1975] is performed on each 3D mesh across subjects in order to eliminate the shape differences caused by translation, rotation and scaling. The algorithm can be summarized as follows.

1. Select one shape to be the approximate mean shape (e.g. the shape of the first subject).

2. Align each shape to the approximate mean shape:

a. Translate each shape so that its centroid is at the origin.

b. Normalize each shape X by a scale factor s=||X||, where X is the coordinate matrix consisting of row vectors of the coordinate of each point after translation.

c. Rotate each shape to align with the newest approximate mean shape. The details of the rotation algorithm can be found in [Ross 2004].

3. Calculate the new approximate mean shape from the aligned shapes. If it is different from the one in step 2, return to step 2.

The difference in the size of the CC has been eliminated in the spatial alignment, so the shape analysis only reveals pure shape difference in local CC area between patients and controls. In the local shape analysis, difference between groups at every surface location is tested. Similar to 4.1.5, Hotelling T^2 two-sample metric is used to test the group difference in the point coordinates (3D variables) and FDR is used as the appropriate approach for p-value correction. Both raw and corrected p values are shown in the experiment.

71

4.2.5 Results

The subjects and MRI data are the same as 4.1.2. Table 4.3 shows the t-test results of the volumes. The mean TBV and ICV of the patients are greater than those of the controls, but the p-values do not reveal any significance in this difference ($p>0.05$). The raw measures of the CC volume and its sub-regions all show that the patient's volume is significantly smaller than the controls ($p<=0.05$), especially in the region r2 (anterior body). The comparison of the scaled measures by TBV and by ICV shows similar results, both of which augment the significance of the difference. Table 4.3 only shows the results normalized by TBV and the results normalized by ICV are similar.

Table 4.4 gives the t-test results of the traits regarding the bounding rectangle and the thickness. In the raw measures, the patients have significantly reduced length and aspect ratio of the bounding rectangle, while other traits do not show any significance in the difference between patients and controls. When scaled by the cubic root of TBV, the length of the bounding rectangle keeps its strong significance and the width still shows no significance in the group difference. The p-values of the difference in thickness are decreased but none of them reach the significance level (0.05). The difference in L2 between patients and controls is close to significant, which is consistent with the volume measurements since the anterior body displays most significant difference. The comparison shows similar results when the measurements are scaled by the cubic root of ICV, and Table 4.4 only shows the results of TBV normalization.

Fig.4.8 gives a descriptive visualization of the group difference. The overlaid average CC structures are shown in Fig.4.8(a). Fig4.8(b) shows the signed distance map between the two average structures rendered on an overall average shape, where the negative distances indicate the patients' structure is outside the controls' and the reverse is for the positive distances. The anterior is more projected and the posterior is more inward in the patients' structure. The posterior body and isthmus are thinner in the patients' structure. The significance maps of the raw and corrected p-values are shown as color coded p-values in Fig. 4.8(c) and (d). Smaller p-values indicate larger statistical significance. A two-tailed alpha level of 0.05 is chosen as the significance threshold. Fig.4.8(c) is the significance map of raw p-values, which reveals significant difference in the anterior most, anterior body, isthmus and posterior bottom. This result is an optimistic estimation. The FDR result is shown in Fig.4.8(d), which retains part of the significance in the raw map.

+   -   p=0   p>=0.05
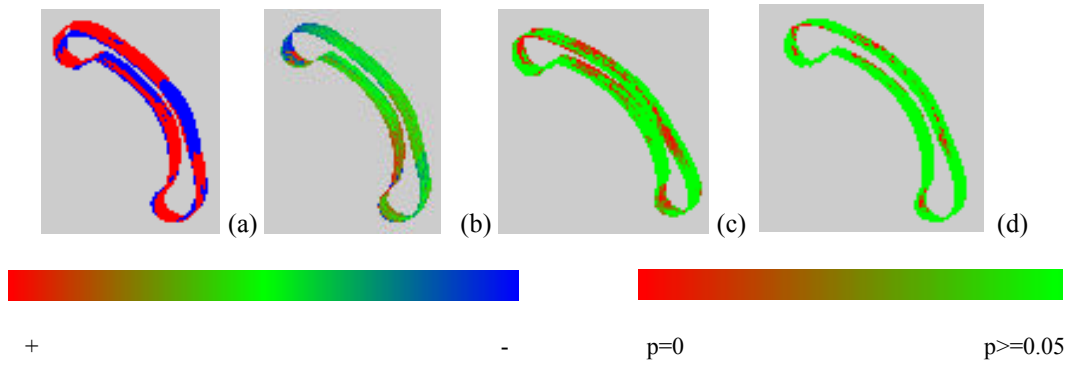
Fig.4.8. Statistical testing results: (a) Overlaid average structures (blue: controls, red: patients) (b) Distance map between the two averages (c) raw significance map (d) corrected significance map

Table 4.3   Results of the volume comparison

CC Measure	Patients	Controls	p-value
TBV (cm^3)	1421.9±299.4	1313.6±336.0	0.09
ICV (cm^3)	1460.4±308.3	1348.3±345.5	0.09
Raw measures	(mm^3)		
Total CC	4229.1±755.3	4727.7.0±971.4	0.01
r1	1754.2±321.3	1939.5±367.9	0.02
r2	474.4±103.9	556.1±142.9	0.008
r3	439.7±123.1	505.1±139.9	0.03
r4	404.6±108.3	458.8±140.8	0.05
r5	1153.3±227.7	1263.5±275.7	0.05
Scaled by TBV	($\times 10^{-4}$)		
Total CC	31.0±9.3	37.0±8.1	0.004
r1	13.0±3.5	15.1±3.1	0.004
r2	3.4±1.0	4.3±1.0	0.001
r3	3.2±1.2	3.9±1.0	0.009
r4	2.9±1.0	3.6±0.9	0.01
r5	8.6±2.9	10.1±2.4	0.02

Table 4.4 Results of the global traits analysis

CC Measure	Patients	Controls	p-value

Raw measures	(mm)		
Length	66.2±4.8	72.1±5.0	$<10^{-4}$
Width	24.0±3.0	23.6±3.6	0.33
Length/width	2.78±0.3	3.11±0.49	0.002
L1	5.91±1.13	6.09±1.19	0.28
L2	5.40±1.18	5.76±1.35	0.14
L3	5.11±1.35	5.29±1.53	0.31
L4	7.53±1.76	7.77±2.06	0.32
Scaled by TBV			
Length	0.59±0.05	0.66±0.06	$<10^{-4}$
Width	0.21±0.02	0.22±0.03	0.35
L1	0.053±0.011	0.056±0.010	0.14
L2	0.048±0.011	0.053±0.011	0.07
L3	0.045±0.012	0.048±0.013	0.20
L4	0.067±0.016	0.071±0.017	0.22

4.2.6 Conclusion

This section demonstrates a successful application of the planar conformal mapping to the statistical shape analysis. The clinical findings of the CC shape analysis are very similar to the results in 4.1, except the additional dimension is added. Besides the shape analysis, complete global traits analysis is performed which is an important procedure in clinical. Global traits analysis and shape analysis are complementary to each other. The former provides global information about the difference between pathological structures and normal ones, while the latter reveals the local difference at specific locations. Although global traits analysis is not the focus of this thesis, it is included in this section for completeness.

4.3 Shape Analysis of Corpus Callosum in Phenylketonuria

4.3.1 Background

Phenylketonuria (PKU) is a metabolic disorder resulting from a deficiency of phenylalanine hydroxylase, which leads to elevated levels of the amino acid phenylalanine (Phe) in the blood. It is an autosomal

recessive disorder caused by mutations in the PAH gene. The clinical behavior of untreated state is characterized by mental retardation, macrocephaly, seizures, eczema, and other behavior abnormalities [National Institutes of Health Consensus Development Panel, 2001], and the brain pathology includes reduced brain weight and hypo- and demyelination of cerebral white matter [Scriver et al., 1995]. However, the heterogeneity in the mutations of PAH results in wide phenotypic heterogeneity, and the relationship between the clinical phenotype and the genotype is not always constant [National Institutes of Health Consensus Development Panel, 2001].

By the application of a Phe-restricted diet beginning in the first days of life, the phenotype of untreated PKU can be prevented. Due to the prevalence of Phe in foods, even under dietary restriction, individuals with PKU still have higher-than-normal levels of Phe in their systems. Many comparisons between early treated PKU patients and controls showed cognitive and behavioral deficits in the patients. MR imaging studies on early treated PKU patients have primarily shown the volume loss in white matter regions of the brain in the patients [Thompson et al., 1993; Phillips et al., 2001; Hasselbalch et al., 1996; Pfaendner et al., 2005; Hahnel, 2008], although gray matter changes in cortical layering, tissue mass (atrophy), and reduced dendritic arborization have also been reported [Dyer et al., 1996; Dyer, 2000]. The corpus callosum has been reported to be one of the most severely affected structures with respect to volume loss in early treated PKU patients [Pfaendner et al., 2005; Hahnel, 2008]. Studies have found that early treated PKU patients displayed a disruption in communication between hemispheres, caused by abnormal myelination of neurons that make up the corpus callosum [Gourovitch et al., 1994; Banich et al., 2000].

Similar to 4.2, we perform global traits analysis as well as the shape analysis of the CC in early treated PKU. However, the CC model is a closed surface obtained from direct 3D segmentation instead of stacking 2d contours. Therefore, we apply the method in 3.2 for the shape correspondence.

4.3.2 Participants and Data

Eight participants with PKU (4 males, 4 females) who were diagnosed at birth and maintained on a phenylalanine-restricted diet participated in the study and ranged in age from 11 to 27 years ($M = 16.80$, $SD = 4.92$). Eight individuals without PKU (4 males, 4 females) comprised an age-, education-, and gender-matched control group. The control group ranged in age from 12 to 27 years ($M = 17.19$, $SD = 5.05$). Individuals with psychiatric and/or medical history unrelated to PKU were excluded.

All MRI Scans were obtained on a 1.5T Siemens Symphony scanner with a standard 8-channel head coil. A high-resolution three-dimensional T1-weighted sagittal scan was collected for purposes of structural analysis (MP-RAGE sequence: TR = 1920 ms, TE = 4 ms, flip angle = 8°, in-plane resolution = 1 x 1 mm, slice thickness = 1 mm, number of slices = 160).

The CC models were obtained using the semi-automatic method in 2.2. To further increase the accuracy and validity of the results, we ensured that all segmentations were completed by the same trained individual using the aforementioned method. An implicit voxel representation (i.e. a 3D binary image) and an explicit mesh representation of the model were both obtained after the segmentation. Each model was a 3D mesh with 3000~4000 points. The segmented CC model can be seen in Fig.4.9.

4.3.3 Global Traits Analysis

We used the voxel representation for the volume analysis. The volume analysis is performed in the same way as in 4.2.3, except the voxel count of the CC can be directly calculated from the 3D model instead of the summation of the area on each slice.

We explored MRI traits based on the oriented bounding box of the CC model. Explicit representation of the model was used for this analysis. Principal component analysis (PCA) [Jolliffe, 1986] was used to find the principal orthogonal axes of the CC model, and the edges of the bounding box were parallel to one of these axes. Fig.4.9 shows two views of the CC model along with the bounding box. The length was defined as the frame which extended from the posterior to the anterior of the CC, the width was the frame that extended perpendicular to the sagittal plane, and the height was the other frame. We also calculated the ratio between each pair of orthogonal frames. The shape of the bounding box depends on both the size and the shape of the CC, so these MRI traits can reflect both size and shape information of the CC.

To account for the brain volume effect, we scaled the above raw measurements by the cubic root of TBV and ICV respectively. Since these measurements were in mm, this scaling made the units consistent. Both raw and scaled measurements were compared between patients and controls.

76
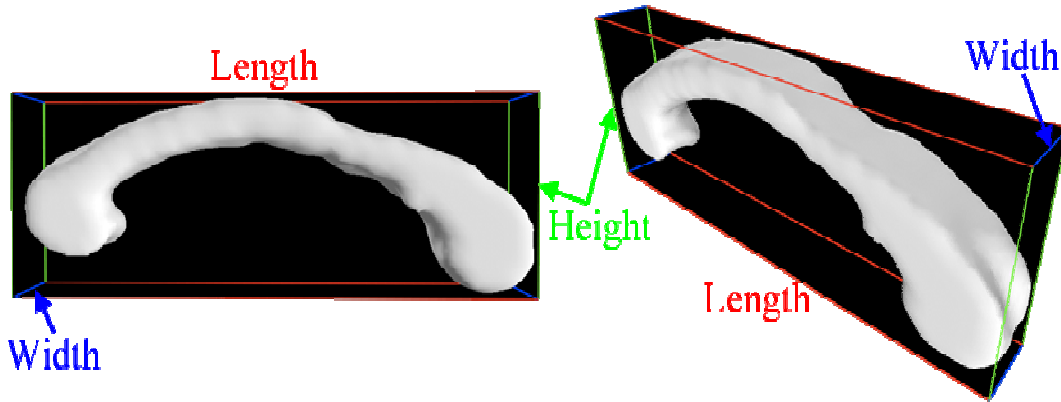
Fig.4.9 Two views of a 3D CC model and its oriented bounding box (red line: the length; green line: height; blue line: width)

4.3.4 Statistical Shape Analysis

Point correspondence among individual shapes is established using the method in 3.2. One of the 16 meshes was randomly selected as the template, and 767 landmarks were sampled on this template. The result was 16 sets of corresponded landmarks, which can be used directly for statistical shape analysis. Since each shape is spatially aligned with the template before the shape correspondence is established, no additional spatial alignment is necessary among all the shapes. The difference in the size of the CC has also been eliminated in the spatial alignment prior to the shape correspondence, so the shape analysis only reveals pure shape difference between patients and controls. Hotelling T^2 two-sample metric is used to measure how two groups of shapes are different from each other at each surface location and FDR is used for p-value correction.

4.3.5 Results

Two-tailed t-tests and two-tailed Mann-Whitney-Wilcoxon (MWW) [Conover, 1980] tests were used for statistical comparison of the MRI traits. The MWW test is similar to t-test and primarily suited for smaller data sets such as ours. A significance level of 0.05 was selected both tests. Table 4.5 shows the results of the two tests in each comparison.

The two testing methods showed similar results. No significant difference was found between the two groups in the brain volumes, raw CC volumes and scaled CC volumes. Both tests indicated significant difference in the height of the bounding box of the CC. Both raw and scaled values of the height were significantly shorter in PKU patients than in the controls.

77

Fig.4.10 gives a descriptive visualization of the group shape difference. Fig.4.10(a) shows the overlaid average shape of each group, and Fig.4(b) shows the signed distance map between the two average shapes rendered on an overall average shape of all 16 subjects. By examining the average structures, we find that in the patients the body of the CC is less bent, the anterior tip (genu) and posterior tip (splenium) are shorter, and the anterior most is more projected. Note that the triangle mesh information is lost after landmark sampling, and we reconstruct the surface using COCONE [Dey and Giesen, 2001] for display purpose.

The significance maps of the raw and corrected p-values are shown in Fig.4.10(c-d). Smaller p-values indicate larger statistical significance. A two-tailed alpha level of 0.05 is chosen as the significance threshold for the raw p-values, and FDR threshold p_α is 0.01 for corrected p-values. Fig.4.10(c) shows the significance map of raw p-values, which is an optimistic estimation. The significance map of the corrected p-values (Fig.4.10(d)) shows that the major significant difference lies in the genu and splenium. This is consistent with the distance map in Fig.4.10(b) which shows the main difference between the averages at the these parts.

| + | - | p=0 | p>=0.05 |
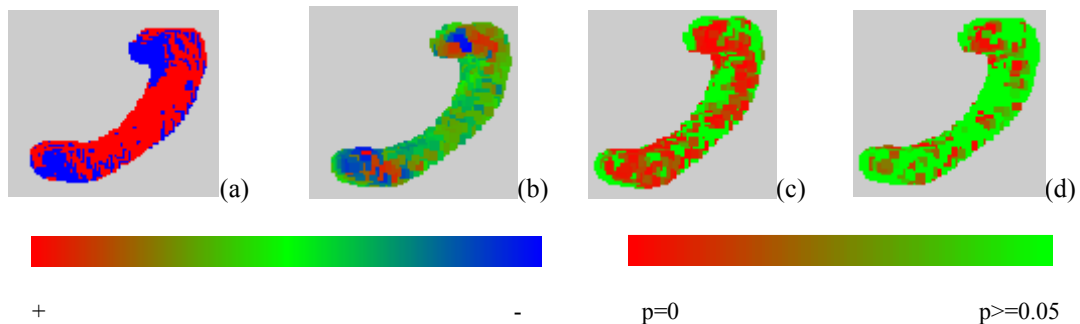
Fig.4. 10 (a) two views of the overlaid average shapes (blue: controls, red: patients) (b) two views of the distance map between the two average shapes (negative distances indicate the patients' structure is inside the controls' and the reverse is for the positive distances) (c) raw significance map (d) corrected significance map

Table 4.5. Results of the CC analysis comparing patients with PKU to control data (Bold entries indicate a significant difference between the patients and the controls).

Measurements	Patients		Controls		t statistic	p-value	U statistic
	Mean	Std	Mean	Std			
TBV (x104 mm^3)	128	8.68	126	7.73	0.2891	0.7767	29

ICV ($\times 10^4$ mm^3)	147	11.7	142	13.2	0.7884	0.4436	26
CC volume ($\times 10^4$ mm^3)	54.2	0.0636	55.3	0.0631	-0.3314	0.7453	31
CC volume/TBV($\times 10^{-3}$)	4.3	0.535	4.4	0.638	-0.4426	0.6648	30
CC volume/ICV ($\times 10^{-3}$)	3.7	0.461	3.9	0.781	-0.7490	0.4663	30
CC length (mm)	71.9	4.47	76.2	6.18	-1.5712	0.1385	21
CC height (mm)	26.9	5.68	33.0	5.43	-2.1848	0.0464	12
CC width (mm)	10.4	1.28	11.0	2.48	-0.6789	0.5083	30
CC length/ CC height	2.78	0.619	2.35	0.219	1.7958	0.0941	15
CC length/ CC width	7.03	0.889	7.19	1.58	-0.2513	0.8053	40
CC height/ CC width	2.63	0.633	3.13	0.903	-1.2727	0.2238	22
CC length/(TBV)$^{1/3}$	0.664	0.0399	0.706	0.0693	-1.4926	0.1577	24
CC height/(TBV)$^{1/3}$	0.247	0.0492	0.306	0.0550	-2.2294	**0.0427**	12
CC width/(TBV)$^{1/3}$	0.956	0.0127	0.102	0.0216	-0.6977	0.4968	28
CC length/(ICV)$^{1/3}$	0.633	0.0362	0.680	0.0729	-1.6268	0.1261	24
CC height/(ICV)$^{1/3}$	0.236	0.0470	0.295	0.0566	-2.2547	**0.0407**	11
CC width/(ICV)$^{1/3}$	0.912	0.0120	0.0979	0.0203	-0.8032	0.4353	28

4.3.6 Discussion and Conclusion

This section demonstrates the direct application of the method in 3.2 to the clinical shape analysis. The point density in the resulting shape is greatly reduced, and the statistical analysis can only be performed on the low resolution surface. However, sufficient details of the shape difference can still be reflected (Fig.4.10), which can meet most clinical needs. Note that the shape analysis in this section is not pseudo-landmark analysis because no anatomical landmark constraints are included. It can be considered as a reduced whole surface analysis since it has the same idea with whole surface analysis except for the reduced point density. As in 4.2, global traits analysis is also shown for completeness of this study. Since the CC model used in this study is 3D in nature, the global traits are designed to represent the features of a 3D surface instead of individual 2D slices.

We found no significant volume difference in the CC between early treated PKU patients and the controls, but we did find reduced CC height and significant shape abnormalities in the genu and splenium of the CC

in PKU patients. Most previous studies reported white matter loss in early treated PKU patients, and the CC is one of the most severely affected structures. These volumetric findings in this study may help to determine whether patients with PKU should be recommended to receive lifelong dietary treatment or not. However, we did not find any significant difference in the raw and scaled CC volumes between PKU and controls. This might be due to our small sample size or the heterogeneity of PKU population. Interestingly, we did find significant reduction in both raw and scaled CC height in PKU patients, which has not been reported in any previous studies. Further more, we find significant shape abnormalities in the genu and splenium of the CC in PKU patients. This local shape morphology cannot be detected using global analysis such as MRI traits analysis, and it also has not been reported before. This might serve as additional evidence for the association with abnormal myelination of neurons that make up the CC. It is suggested that the relationship between variations in the phenotypes associated with blood Phe concentration be studied [National Institutes of Health Consensus Development Panel, 2001]. Correlation between volumes of brain structures and blood Phe value has been studied and no significant correlation was found [Pfaendner et al., 2005; Hahnel, 2008]. Other global traits can serve as additional phenotypes to the volume, and their relation with blood Phe value can be studied in a similar manner. This will provide stronger evidence of whether or not patients with PKU should receive lifelong dietary treatment.

The clinical results of this study need to be interpreted cautiously due to the limitation of the sample size. The sample size is small compared with other statistical analysis, so the results may not be representative enough for large population. Since previous brain MRI studies revealed only volume loss of the CC in PKU patients and few have conducted local shape analysis or other MRI traits analysis of the CC, we cannot compare our results with other works. Therefore, this study serves as a pilot study about the shape morphology of the CC in early treated PKU and further study with larger sample sizes is needed to confirm the results.

4.4 Gender Difference of Lateral Ventricles

4.4.1 Background

The ventricular system is a structure in the center of the brain filled with fluid. Structural changes in the lateral ventricles are related to various neurodegenerative diseases. For example, [Ferrarinia et al., 2006,

Thompson et al., 2004] found the lateral horns of the ventricles are different between Alzheimer's patients and normal people, [Styner et al., 2005] studied the genetic influence of the lateral ventricle shapes in schizophrenia, and [Chou et al., 2007] reported enlarged ventricles in autistic patients. Gender differences of the ventricles have also been studied on normal subjects or subjects with certain neurological disorder. [Giedd et al., 1997] found more prominent growing of the lateral ventricle size in male than in female between 4 and 18 years of age. [Nopoulos et al., 1997] found greater enlargement of the ventricles in male schizophrenia patients than in female. [Yotsutsuji et al., 2003] found that schizophrenia patients showed ventricular enlargement, particularly in the left temporal horn, being more severely affected in the male than in the female. Studying sexual dimorphism of brain structures among control group can provide a reference for studies on brain abnormalities of certain neurodevelopmental disorders, especially gender-dominant disorders such as autism (male-female ratio is 4:1). The aim of this paper is to study the shape differences of the lateral ventricles between normal male and female subjects.

Anatomical landmarks are first identified and pseudo-landmark analysis is performed. The method for shape correspondence is similar to 4.1.5. However, the case for 3D shapes is more complicated and a direct application of the method in 4.1.5 to 3D shapes is not feasible. We incorporate the method in 3.2 with anatomical landmark constraints to perform the pseudo-landmark analysis.

4.4.2 Subjects and Data

14 female and 10 male normal subjects were recruited for this study. No subject had been mentally ill or had any first degree relatives with a psychotic disorder. Student t-test was used to compare the mean age between genders and no significant difference was found. Table 4.6 shows the age distribution of the participants. This study was approved by the local Institutional Review Board.

Brain MRI scans of the subjects were obtained on a 3T Siemens Trio scanner with a standard 8-channel head coil at the Brain Imaging Center in the Department of Psychological Sciences at the University of Missouri. A custom VacFix® Cushion (S&S Par Scientific Inc.) were used to provide head and upper body restraint. To aid in alignment of the subsequent structural scan, a set of structural scout images were collected first using a standard T1-weighted pulse sequence. Next, a high-resolution three-dimensional T1-weighted sagittal scan was collected (MP-RAGE sequence: TR = 1920 ms, TE = 3.75 ms, flip angle = 8°, in-plane resolution = 1 x 1 mm, slice thickness = 1 mm, number of images = 160, matrix =256 x 256). The

data were aligned with the stereotactical coordinate system and re-sliced to isotropic voxels of 1mm3 using Slicer (www.slicer.org).

The 3D ventricle models were obtained using the semiautomatic segmentation method in 2.3. All segmentations were completed by the same operator, and each result was inspected and adjusted by an expert to ensure the adequacy of the model. Each ventricle model included ventricular body, occipital horn and lateral horn, and the temporal horn was excluded from the segmentation result. The segmentation result was a signed distance field and was converted to a triangle mesh using marching cube algorithm [Lorensen and Cline, 1987]. Fig.4.11 shows some ventricle models from different subjects. Each model was a 3D mesh with 6000~7000 points.

Table 4.6. Age statistics of the participants

Age	Male	Female	P value
Mean	23.72	24.88	0.56
Max	27.90	31.35	
Min	20.65	20.47	

(a)   (b)
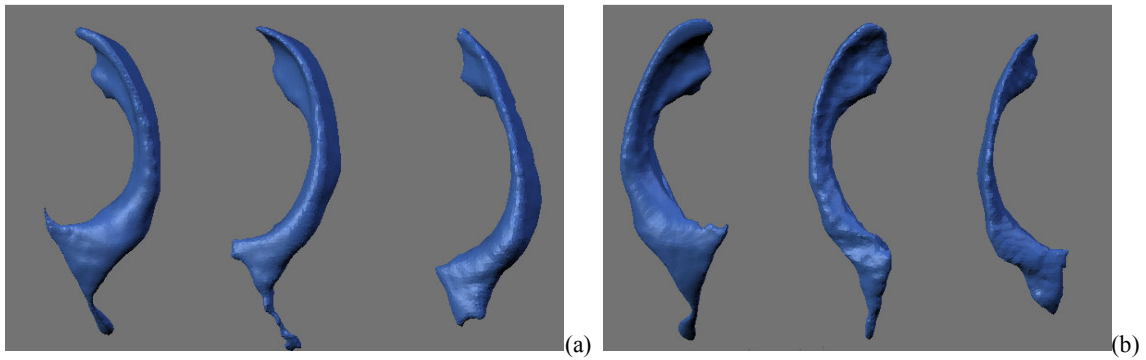
Fig.4.11. 3D segmentation results of (a) left and (b) right ventricles

4.4.3 Shape Correspondence with Landmark Constraints

We define six landmarks that are anatomically meaningful and geometrically prominent (Fig.4.12). These landmarks are assumed to be corresponded across difference shape instances because of their common geometrical and anatomical features. The advantage of identifying anatomical landmarks is that there is exact matching between landmarks on different shape instances. This matching is usually more crucial in the subsequent shape comparison. The landmarks were identified using a semiautomatic method. The user first clicked each landmark on the 3D view of the ventricle model. A program then searched around a small

neighborhood of the clicked point automatically and found the maximum curvature point which will be used as the landmark. The landmark selection for all the data was operated by a trained expert.

In order to find the correspondence of the rest points on the shape, a modified version of the method in 3.2 is applied. As in 3.2, a template shape was randomly selected from all the shape instances and a subgroup of points are uniformly sampled using 3D grid cells. The purpose of this sampling is to reduce the computational cost in the later shape analysis, and only the sampled points are used to establish the 3D correspondence. The size of the cell controls the density of the sampled points, and in our implementation about 2000 points are retained compared to 7000 points on the original model (Fig.4.13(a)). During the sampling the six anatomical landmarks were not allowed to be removed.

The sampled template shape was then transformed to each of the rest shape instances (target shapes) by TPS with the six landmarks as the control points. The theory of TPS can be found in 3.2.3. The control points of TPS in 3.2.3 are the entire set of the sampled surface points, while in this section only the six anatomical landmarks serve as the control points (n=6). The estimated transformation is then applied to all the sampled points on the template. After TPS transformation, the landmarks on the template are moved to the exact locations of the corresponding landmarks on the target shape, so the template and the target shapes are aligned through the landmarks (Fig.4.13(b)). The correspondence of the remaining points between the transformed template and the target shape is constructed using the method in 3.2.2. The result is a subset of the target points which are corresponded to the sampled points on the template (Fig.4.13(c)). Note that the six landmarks were already corresponded between the two shapes, so they were excluded in the process of finding the initial correspondence.
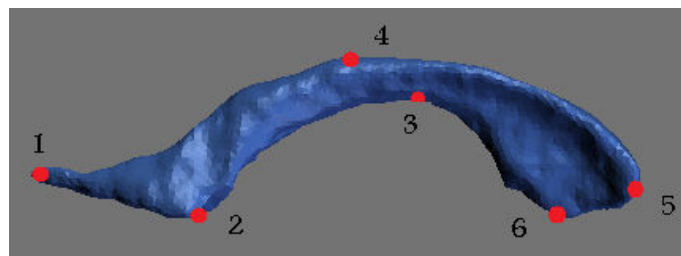
Fig.4.12 Anatomical landmarks for the lateral ventricle
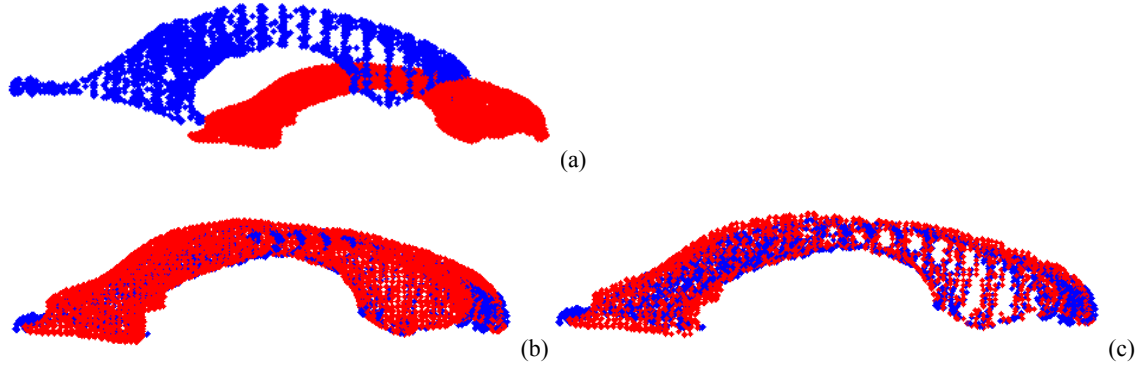
Fig. 4.13 (a) sampled template shape (blue) and original target shape (red) (b) sampled template shape after TPS transformation and the target shape (c) sampled template shape after TPS transformation and the corresponding points on the target shape (only point clouds are shown without triangle meshes)

4.4.4 Statistical Shape Analysis

Generalized Procrustes Analysis is performed on all the 3D shapes in order to eliminate the shape differences caused by translation, rotation and scaling. In the local shape analysis, difference between groups at every surface location is tested. Hotelling T^2 two-sample metric is used to measure how two groups are locally different from each other and FDR for p-value correction. FDR method allows the false positive to be within a small proportion α (α =0.01 in our experiment).

4.4.5 Results

Fig.4.14 shows the shape comparison results of the left and right ventricles. The shapes are shown as point clouds. The average shapes of the two groups are overlaid to give a descriptive visualization. From the overlaid structures, we find that both left and right ventricles of the females are more projective in the anterior horn and smaller in the posterior horn. A two-tailed alpha level of 0.05 is chosen as the significance threshold for the raw p-values. The p-values at each surface point are shown as a significance map on an overall mean shape. The raw significance map of the right ventricle shows significant shape difference in part of the anterior horn between male and female, but the corrected significance map shows no shape difference between the two groups. The raw significance map of the left ventricle shows significant shape difference in the anterior horn and posterior body between male and female, but the corrected significance map only retains the difference in part of the anterior horn between the two groups.
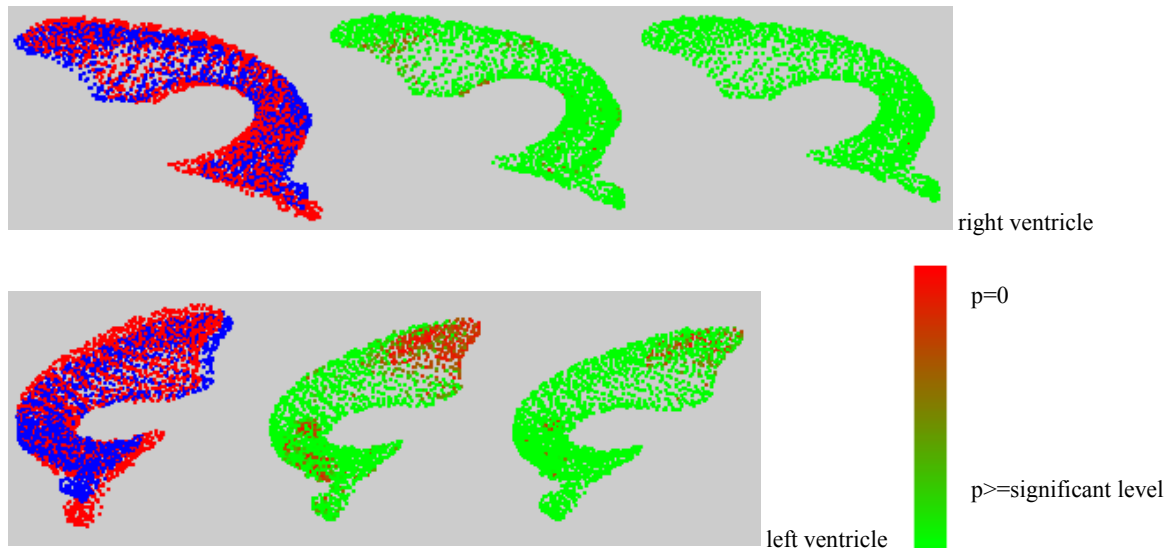
Fig.4.14. On each row: left: average ventricle shapes of females (blue) and males (red) overlaid; middle: significance map of raw p-values; right: significance map of p-values after FDR correction

4.4.6 Conclusion

This section studies the shape differences of the lateral ventricles between genders through pseudo-landmark analysis. The shape correspondence method in 3.2 is applied with some modifications to accommodate the constraints of anatomical landmarks. As stated in 1.3, the advantage of pseudo-landmark analysis is that anatomical landmarks are exactly corresponded to the same anatomical landmarks among different shapes. For anatomical shapes such as brain structures, there are usually obvious landmark points that should be exactly matched among different shapes. Mismatch of landmark points can lead to more serious problems in the shape analysis than mismatch of non-landmark points. Therefore, it is crucial to ensure the correspondence of landmarks on anatomical shapes. We enforce a semiautomatic landmark selection step before establishing the correspondence of other surface points. This additional step seems trivial, but it ensures the correspondence of landmarks which is an essential requirement in anatomical shape analysis. Note that any previous methods of shape correspondence can be incorporated with the anatomical landmark constraints with varying complexity, but the method in 3.2 is particularly suitable for pseudo-landmark analysis, which has been demonstrated in 4.4.3.

Significant difference in the anterior horn of the right ventricle between males and females is found in this study, which had not been reported before. Most previous studies only focused on volume differences. As there is no previous work on this analysis, we cannot compare our results with others. The abnormalities of

85

the ventricles have been studied most in schizophrenia and Alzheimer disease. The results of our study can be used as reference for disease related studies on ventricle shapes. For example, it may explain whether the gender difference of lateral ventricles in a certain disorder is disorder specific or due to intrinsic gender difference. Due to the small sample size, this study serves as a pilot study about the shape morphology of the ventricles between genders and further study with larger sample sizes is needed to confirm the results.

4.5 Summary

This chapter gives several real examples of statistical shape analysis of brain structures in clinical cases. The segmentation methods in chapter 2 and shape correspondence methods in chapter 3 are applied here. Global traits analysis is also performed in some sections as part of the overall analysis procedure. This chapter completes the shape modelling framework of brain image analysis. A flowchart of this framework is summarized in Fig.4.15 and detailed methods are presented in previous chapters. Since our focus is on brain structures, some of the proposed segmentation methods are particularly targeted to certain brain structures. However, the proposed shape correspondence methods can be applied to generic shapes, and the entire framework can be generalized to other medical images and other clinical cases.
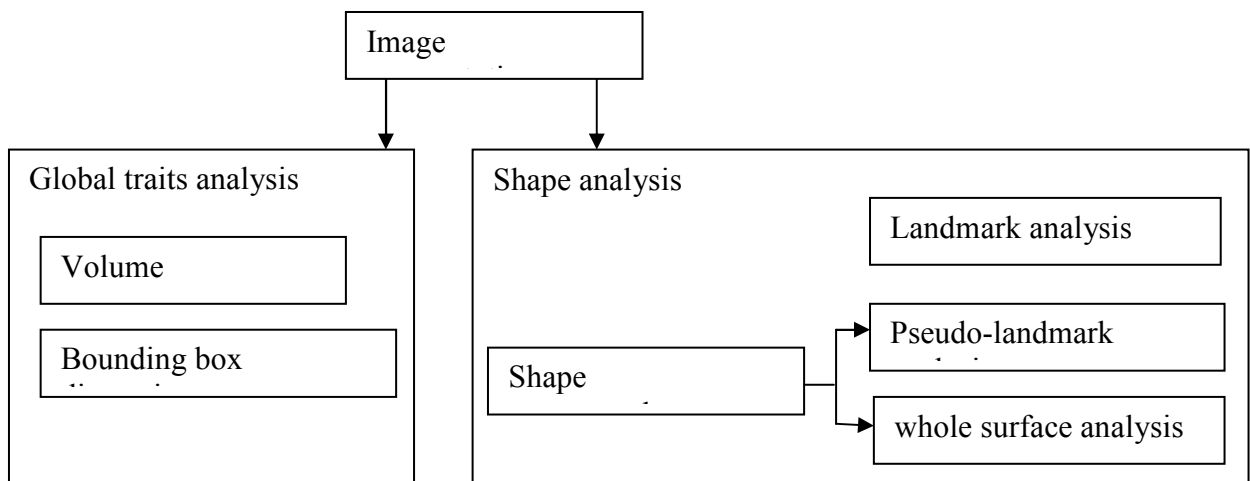
Fig.4.15 Framework of medical image analysis

5. Automatic Detailed Localization of Facial Features

5.1 Background

Automatic localization of facial features is important technique in many applications, such as face recognition, human-computer interaction, surveillance, etc. Particularly, computer aided diagnosis is becoming an attracting application in recent years. For example, the presence of a specific facial pattern in patients with a genetic syndrome indicates phenotypic expression of this disease, and the question has been asked whether a computer can recognize and classify dysmorphic faces solely on the basis of 2D images. [Dalal and Phadk, 2007; Loos et al., 2003; Boehringer et al., 2006] performed syndrome classification based on the feature points extracted from facial images, but the feature points were manually identified in their works. We aim at developing an automatic method to locate facial features for the purpose of computer based diagnosis, but the approach is applicable to other fields with similar requirement.

Various methods for facial feature detection/localization have been reported in literatures. [Feris et al., 2001] proposed a two-level Garbor wavelet network (GWN) to detect eight facial features. In [Bhuiyan et al., 2003] six facial features are detected. Face detection is performed using skin color segmentation and genetic algorithm. Image enhancement is then performed and the facial features are detected through intensity threshold. In [Gourier et al., 2004], an elliptic facial region is detected based on skin color. A vector of Gaussian derivatives and the linear combination of these descriptors act as detection functions for facial features. In [Cristinacce et al., 2004], a multi-stage approach is used to locate 17 facial features. The face is detected using boosted cascaded classifier [Viola and Jones 2004], and the same classifier is trained using facial feature patches. Further refinement is achieved by active appearance model [Cootes et al. 2001]. [Asteriadis et al., 2009] detect eye centers and mouth corners using a hierarchical scheme. After face detection and normalization in size, a distance vector field (DVF) is created according to the distance from each pixel to the closest edge. The eye and mouth regions are found by matching the DVF with a template. Intensity and color information is then used to locate eye centers and mouth corners.

 All the above works only look for feature points. However, more details of facial features lie in their shape information, such as eyelid and lip contours. While it may not be too tedious to label a few feature points

manually, it is almost impractical to delineate shapes of facial features by hand. Therefore, an automatic procedure for facial feature contour extraction is in demand, which is the main focus of this paper.

Lip contour extraction has been a branch with extensive studies in the past decade. Since the lip is rich in color, most previous works utilized the color information to separate the lip form the background, and then some refined methods were used to extract the accurate lip contour. [Wang et al., 2004] used three quadratic curves to characterize the lip contour, with 16 points on the curves as parameters. The lip region was first separated from the background using fuzzy C-means clustering with shape constraints [Wang et al., 2002]. [Eveno et al., 2004] used several cubic curves to model the lip contour. Feature points on the upper lip were first detected using "jumping snake" which is robust to the initialization. Hybrid edge [Eveno et al., 2002] was computed from pseudo-hue which intensifies the upper lip edge. [Yokogawa et al., 2007] matched a template of four quadratic curves to the lip contour, and the lip was first separated from the background by thresholding the hue component. Active contour model was also used in several literatures. [Wakasugi et al., 2004] proposed an edge detection method based on the separability of two color intensity distributions, which was used as the image energy in the active contour model. A level set method without re-initialization was proposed in [Sohail and Bhattacharya, 2007], where only grayscale images were used.

Much fewer studies exist for eyelid and nose contour extraction. [Vezhnevets and Degtiareva, 2003] proposed an eye contour extraction method. Iris was first detected based on its intensity and circular shape. The points on the upper eyelid were then detected by Hough transform. A cubic curve was used to fit the upper eyelid points and a quadratic curve was used to fit the lower eyelid. [Zheng et al., 2005] used Gabor filter to detect the eye corners and the top and bottom points of the iris, which were then used as control points for a spline curve to fit each eyelid. Earlier works [Kampmann and Zhang, 1998; Yuille et al., 1992] used edge based deformable models to fit the shape contours of facial features, but it has been stated in [Vezhnevets and Degtiareva, 2003] that the edge information is not robust due to spurious edges and discontinuous edges. [Ding and Martinez, 2008] performed detailed facial feature localization. This is the most complete framework in that it locates the shape of the eyes, nose, mouth, eyebrows and the lower chin. However, the main contribution of [Ding and Martinez, 2008] is the facial component detection based on subclass discriminant analysis and a context sensitive classifier design, thus the boundary localization of

each facial component is not the focus of their work. Most of their localization methods are oversimplified and are not robust enough for varying subjects and image conditions.

We propose a complete framework for automatic detailed facial feature localization (Fig.5.1). Features of the eyes, the nose, the mouth, and the chin are of interest (Fig.5.2), and more explanations of these features will be shown later. Both feature contours and feature points are located. Unlike the aforementioned works on facial feature point detection, we do not detect feature points directly from the image. Instead, feature contours are located first, and most feature points can be found as distinct points on the contours. We believe that locating a feature point in an image is a harder task than locating a set of points on a feature contour, because a group of points on a contour contain much richer information than a single point to distinguish themselves from the rest of the image. That is why feature point detection often involves extensive training on a large sample but feature contours can sometimes be located by a few image processing steps. Our methods do not require any training, and the feature point localization is reduced to finding extreme-value points on a contour.

Our framework is similar to the framework in [Ding and Martinez, 2008]. However, our focus is the localization of the features of each facial component, rather than the detection of each facial component. Therefore, our feature localization methods take into account more details of each feature contour and are not oversimplified as in [Ding and Martinez, 2008].

The primary application of our work is computer based diagnosis, so a few assumptions are made on the input image. We assume the input is a frontal color face image with neutral expression. The face should take the majority of the image area, and should not be occluded by hair, beard, etc. There should be no in-plane rotation or out-of-the-plane rotation of the face. Usually these assumptions are easy to satisfy in clinical applications.
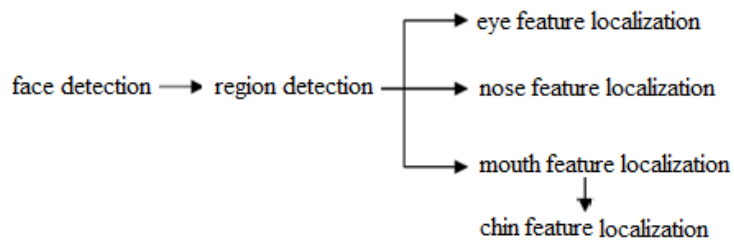
Fig.5.1 Framework of the facial feature localization

(a)   (b)   (c)   (d)
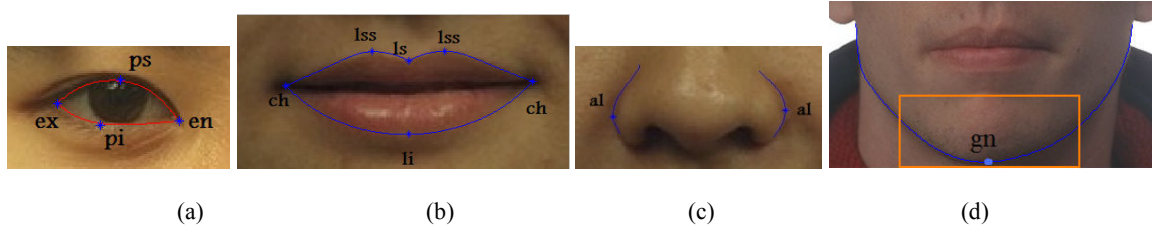
Fig.5.2 Feature points and boundary contours of the (a) eye (b) mouth (c) nose; (d) chin and the bottom chin bounding box

5.2 Face Detection and Region Detection

5.2.1 Face Detection

There are many existing algorithms for face detection, and the details are beyond the scope of this research. Here we apply the widely used Viola-Jones face detector [Viola and Jones, 2001]. The input image is first converted to grayscale for face detection and the subsequent feature detections, except that the color version is retained for mouth feature detection. The image is then cropped to the detected face area as shown in Fig.5.3(b). To facilitate the following region detection, the cropped image is rescaled to a standard size, e.g., the size of the template face image.

5.2.2 Parametric Template Matching

Template matching is the process of comparing a standard pattern (template) with an image for the purpose of finding the occurrence of the template pattern within the search image. The search is performed by moving the template pixel by pixel over the search image while calculating the correlation between the template and the portion of the search image at that position to find the position with the highest normalized correlation value. High computational cost has been a major difficulty of this method. However, if the search area is small this problem can be greatly alleviated. The parametric template [Tanaka et al., 2000] is a linear combination of a set of templates with different feature values (scales, rotations, etc). It has many advantages over traditional template matching in that it can represent geometrical and non-geometrical changes of an object in the parametric template space.

Here we briefly introduce the concept of parametric template. A root template t_0 is defined as the template directly obtained from a template image. A set of vertex templates can be generated from the root template

90

by geometric or non-geometric transformation. Let $t_1, ... t_M$ be M normalized vertex template with the same size, the parametric template is defined as:

$$t_w = (\sum_{i=0}^{M} w_i t_i) / \| \sum_{i=0}^{M} w_i t_i \| \quad (0 \le w_i \le 1, \sum_{i=0}^{M} w_i = 1) \quad (5.1)$$

When comparing the template and an image patch g (normalized) with the same size of the template, our goal is to maximize the normalized correlation $\rho(g, t_w)$ by varying the parameters $\vec{w}$ subject to the constraints in (5.1), which can be solved by Lagrange multiplier method. The solution is

$$\vec{w} = H^{-1}\vec{G} / (\vec{n}, H^{-1}\vec{G})$$

$$\vec{w} = \begin{bmatrix} w_0 \\ \vdots \\ w_M \end{bmatrix}, H = \begin{bmatrix} \rho(t_0, t_0), \cdots, \rho(t_0, t_M) \\ \vdots \quad \ddots \quad \vdots \\ \rho(t_M, t_0), \cdots, \rho(t_M, t_M) \end{bmatrix}, \vec{G} = \begin{bmatrix} \rho(g, t_0) \\ \vdots \\ \rho(g, t_M) \end{bmatrix}, \vec{n} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rho(g, t) = \sum_{x,y} g(x, y) \times t(x, y)$$

$$(5.2)$$

In order to satisfy the constraints of $\vec{w}$ in (5.1), normalization needs to be performed on $\vec{w}$ as follows.

$$\vec{w} = \vec{w} / (\sum_{i=0}^{M} w_i) \quad (5.3)$$

A feature value v_i is bound to each vertex template t_i. After the parameters $\vec{w}$ are obtained, the feature value v of the parametric template t_w can be calculated according to equation (5.4). If v_i is the size of each vertex template, then v is the size of the resulting parametric template. Other features represented by v_i can be interpreted similarly.

$$v = \sum_{i=0}^{M} w_i v_i \quad (5.4)$$

5.2.3 Region Detection

The purpose of region detection is to locate a properly sized bounding box of each facial component. The parametric template method is applied here. Scale variation of the objects is the major concern. Although the face is normalized to a standard size, there is still variation in the size of each individual facial component among different people. We take into account the size variation in both x and y directions. Therefore, a feature vector $\vec{v}_i = [x_i, y_i]$ instead of a scalar is assigned to each vertex template, which consists of the scales in x and y directions relative to the root template.

The root template for each facial component is shown in Fig.5.3(c). Each template is constructed such that some margin is included but no part of other component is included. Note that the chin bounding box is not to be detected in this step. The location of the chin will be determined based on the mouth bounding box, which will be shown in section 5.6. Five vertex templates are generated from each root template, which are scaled in x direction only, y direction only, and both x, y directions. In order to compute the linear combination of the templates according to (5.1), the root template and its derived templates are first aligned by their centers, and all the templates are padded to the size of the largest vertex template so that they have the same size. The initial parametric template has the same size as the largest vertex template.

In the search process, we utilize the prior knowledge about the approximate positions of facial components. Thus, we search for the eyes only in the upper half of the face area, and the right eye and the left eye are searched within the right and left half of the region respectively. The mouth is searched only in the lower half of the face. After the eyes and the mouth are detected, the nose is searched within the region below the eyes and above the mouth.

With the regions detected, the feature localization will be performed on the image patch of each specific region. For the eyes and the nose, the region image patches are cropped from the grayscale image, and for the mouth the region image patch is cropped from the color image.
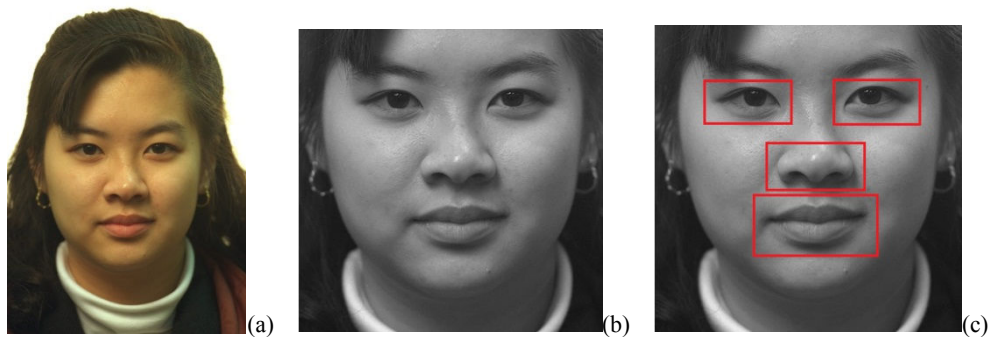
Fig.5.3 (a) the input color image (b) the image converted to grayscale and cropped to the detected face (c) template of each facial component

5.3 Eye Features

We are interested in four feature points shown in Fig.5.2(a) as well as the eyelid contours. Endocanthion (en) is the point at the inner commissure of the eye fissure, ectocanthion (ex) is the point at the outer

commissure of the eye fissure, palpebrale inferius (pi) is the lowest point in the mid-portion of the free margin of the lower eyelid, and palpebrale superius (ps) is the highest point in the mid-portion of the free margin of the upper eyelid. These four points are usually of most interest to clinicians. In the following we describe the method for the right eye, and for the other eye we mirror the eye image first before the feature detection.

5.3.1 Iris Detection

The eye image is first smoothed by a Gaussian filter and normalized in intensity. The reflection in iris is then removed using the method in [Asteriadis et al., 2009]. Briefly, a binary image is generated using otsu's histogram algorithm (Otsu 1979), and any white spot whose area is smaller than 20 pixels is filled with black (Fig.5.4(d)). After that, the corresponding pixels in the original image are filled with the intensity of their surrounding pixels (Fig.5.4(f)).

Template matching method described in 5.2.2 is then used to locate the iris on the binary image, and the root template is a binary image of the iris. The detected iris region is shown in Fig.5.4(e). The top and bottom boundaries are extracted from the detected iris binary image. The top boundary will serve as part of the upper eyelid contour, which will be discussed in the next section. Two key points are located on the bottom boundary, which are the corners of the iris and the sclera (Fig.5.4(f)). They can be found by searching for the first curvature extrema from each side of the bottom boundary.

5.3.2 Eyelid Localization

The eyelid model consists of an upper eyelid curve and a lower eyelid curve, connecting at corner points en and ex. As stated in [Vezhnevets and Degtiareva, 2003], edge map does not give a robust estimation of the eyelids, and luminance valley points along the horizontal lines were detected as the candidate upper eyelid points. Our method is mostly inspired by [Vezhnevets and Degtiareva, 2003]. To detect the upper eyelid, a local minimum map is constructed where each local minimum point is the luminance valley point in either horizontal or vertical direction (Fig.5.5(a)). Since the middle part of the eyelid contour is already detected in 5.3.1, we only look for the upper eyelid points to the two sides of the iris. As shown in Fig.5.5(a), these two parts of the upper eyelid are well captured by the local minimum map, and two straight lines can approximately fit the local minimum points on these two parts. We apply Hough transform [Duda and Hart, 1972] to detect the two straight lines from the local minimum map. Hough transform is known to be robust

to imperfect data and noise. In order to reduce the search space, we restrict the locations of lines based on the iris detection result in 5.3.1. Particularly, the two key points in Fig.5.4(f) serve as the left and right bound for the right and left parts respectively, and a horizontal line passing the topmost point of the middle upper eyelid serves as the upper bound (Fig.5.5(b)). Hough transform produces a set of lines within each search range, and the line with maximum number of local minimum points is selected. The points on the selected lines are the candidate eyelid points (Fig.5.5(c)).

Some false positive points that exceed the bound of the eye may be included in the detected lines. A corner localization procedure is applied to locate the eye corners from the candidate eyelid points. Only the eyelid points that are in between the two eye corners are retained, and others are discarded (Fig.5.5(d)). The template matching method in 5.2.2 is again applied for eye corner detection. The root templates here are the 9x9 patches centered at the left and right eye corners respectively. Instead of searching the entire eye image, we only go through the locations of each candidate eyelid point, and the point location at which there is a best match with the template is selected as the corner position.

The retained points on the two sides of the upper eyelid together with the points on the middle eyelid curve detected in 5.3.1 constitute the upper eyelid points. A Bezier curve is fitted to all these points to obtain the upper eyelid contour (Fig.5.5(e)). Since the Bezier curve interpolates the end points, the resulting eyelid contour always pass through the two eye corner points.

Fig.5.5(a) shows that the local minimum map does not capture the lower eyelid very well, because the contrast on the lower eyelid is not as strong. In [Vezhnevets and Degtiareva, 2003] and [Zheng et al., 2005], a quadratic curve is fitted to the eye corners and the lowest point of the iris to generate the lower eyelid. However, this curve may not approximate the lower eyelid very well because a major portion of the lower eyelid in between the three points is missing. We approximate the lower eyelid by a piecewise linear curve which consists of four line segments. They are divided by the following five points: two eye corners (A,E), lower left bound of the iris (B), lower right bound of the iris (D), the lowest point in the middle of the iris (C) (Fig.5.6(a)). The two eye corners have been detected before so they are fixed, but the other three points are flexible to move. We start from the line segment connecting the left corner (A) and the left lower bound of the iris (B). By moving point B along vertical direction within a small range we can form several candidate lines. On each line the average gradient magnitude is calculated and the line with the maximum

gradient magnitude is selected. The same method is used for the line segment between the right corner (E) and the right lower bound of the iris (D). Then we form several polyline candidates connecting the lower left (B), low middle (C), and lower right bound (D) of the iris, by moving the low middle point of the iris along the vertical direction. The polyline with maximum average gradient is selected. Fig.5.6(b) shows the points on the selected line segments for the lower eyelid.

A Bezier curve is fitted to all the points on the above line segments to obtain the lower eyelid contour (Fig.5.6 (c)). The same as the upper eyelid contour, the lower eyelid contour also pass the two eye corner points. Therefore, the upper and lower eyelid contours connected by the same end points form an entire eye boundary (Fig.5.6(d)).

With the upper and lower eyelid contour, The feature points in Fig.2(a) are easy to obtain. En and ex are already detected which are the two eye corners. pi and ps are the lowest and highest points on the eye boundary (Fig.5.6(d)).

(a) (b) (c) (d) (e) (f)
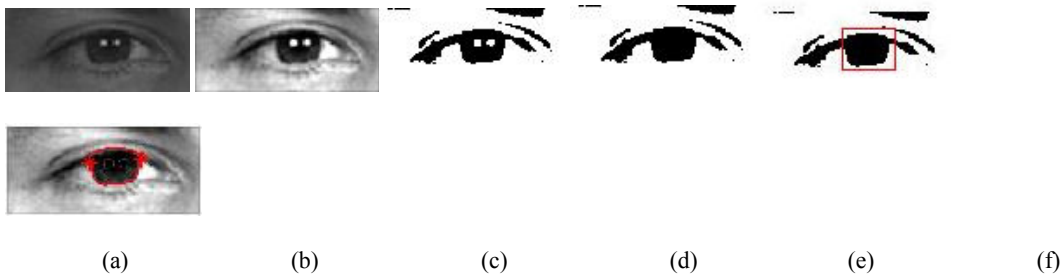
Fig.5.4 (a) the original eye image (b) the result of Gaussian smoothing and intensity normalization of (a) (c) the binary image of (b) (d) the binary image after hole filling (e) the detected iris region (f) boundaries and key points extracted from the iris region, shown on the eye image after reflection removal

(a) (b) (c) (d) (e)
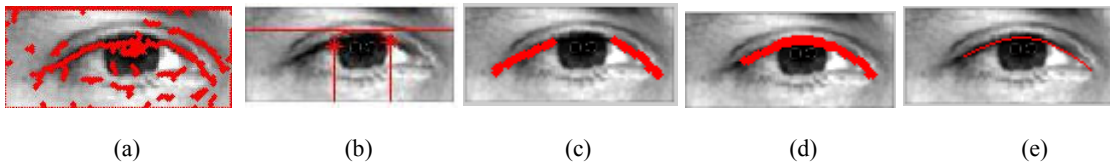
Fig.5.5 (a) local minimum points overlaid on the eye image (b) search range of the upper eyelid points (the lower left and lower right rectangle regions) (c) detected upper eyelid points by Hough transform (d) the entire set of upper eyelid points after eye corner detection (e) final upper eyelid contour

(a) (b) (c) (d)
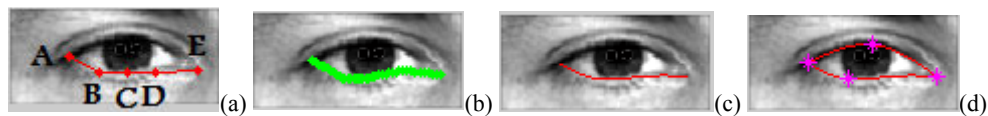
Fig.5.6 (a) piecewise linear curve of the lower eyelid (b) detected lower eyelid points (c) final lower eyelid contour (d) the entire eyelid boundary and the feature points

5.4 Mouth Features

Six feature points (Fig.5.2(b)) are of interest as well as the lip contour. Cheilion (ch) is the mouth corner, labiale inferius (li) is the middle point of the lower lip, labiale superius (ls) is the midpoint of the upper lip, and the projecting point on each side of ls is marked as lss. These points are not only defined from clinical point of view, but also serve as parameters for the curve fitting which will be shown later.

5.4.1 Region Based Lip Segmentation

The mouth is rich in color so color images are often used for mouth feature detection. Various color spaces (HSV, Lab, Luv, RGB) have been explored [e.g.,Wang et al., 2004; Eveno et al., 2004; Yokogawa et al., 2007]. A fuzzy C-means (FCM) method incorporating shape information [Wang et al., 2002] has been proposed to cluster the lip region based on selected color components. This method assumes an ellipse shape of the mouth and can better eliminate noisy regions produced from FCM. However, when the mouth is long and thin the ellipse model cannot approximate it very well. A new color transformation [Eveno et al., 2001] represents each pixel as the curvature of a parabola defined by three points in the transformed color space. Generally this method can generate regions close to the mouth shape, but a threshold needs to be selected manually for different images. In [Lievin et al., 1999], the ratio image R/G is computed and a threshold is automatically selected based on the histogram of this ratio image. This threshold can lead to a binary image that clear separates the mouth and the skin, but when the mouth pixels and the skin pixels are not separable in the histogram of the ratio image, the generated binary image will be erroneous. [Lucey et al., 2002] compared the ratio image R/G and two other color transformations for mouth segmentation. The performance of each method varies depending on the characteristics of the image, and the average performance of each method is about the same. Moreover, [Lucey et al., 2002] points out that for some images none of these color transformations can separate the mouth from the skin, which is an intrinsic problem of chromatic segmentation and more complex approaches are in demand for images with low mouth/skin contrast.

By examining many color based methods for mouth segmentation, we find that the color transform proposed by [Canzlerm and Dziurzyk, 2002] is most robust to low mouth/skin contrast. It transforms an RGB color image (Fig.5.7(a)) to a grayscale image (Fig.5.7(b)) according to equation (5.5).

$$I=(2G-R-0.5B)/4 \quad (5.5)$$

This is a contrast enhanced grayscale image, and a binary image can be easily generated using Otsu's histogram (Fig.5.7(c)). Connected components algorithm [Ronsen 1984] is performed to remove small background noise and holes in the mouth region. Then the boundary of the mouth is extracted to get an initial contour (Fig.5.7(d)).
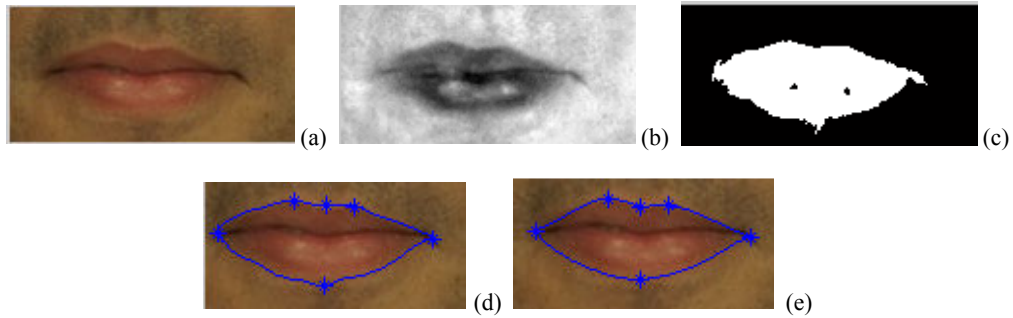
Fig.5.7 (a) the original mouth image (b) contrast enhanced grayscale image (c) the binary image (d) the mouth contour extracted from (c) (e) the final mouth contour and the feature points

5.4.2 Parabola Model for the Lip Contour

We use parabolas to model the lip contour. Four parabolic curves (*yl, ylr, yrl, yr*) are used to approximate the upper lip and one (*yb*) to approximate the lower lip (Fig.5.8). The intersection points of the upper lip curves are $\vec{p}_1$, $\vec{p}_2$ and $\vec{p}_3$, where $\vec{p}_1$ and $\vec{p}_3$ are also the zero-derivative points of the connecting parabolas. $\vec{p}_4$ is the zero-derivative point of curve *yb*. The reason to use four parabolas instead of two for the upper lip is that the curve between $\vec{p}_1$ and $\vec{p}_3$ may not follow the same parametric form as the curve on the other side of $\vec{p}_1$ and $\vec{p}_3$. Moreover, bending and expansion are applied to the point (x, y) on the curve as follows.

$$x' = T(B(x, y), y) \qquad (5.6)$$

where x' is the new x coordinate of point (x, y) and the y coordinate does not change. $B(x, y)$ is the bending function [Wö rz and Rohr, 2006] and $T(x, y)$ is the expansion function defined as follows:

$$B(x, y) = x - y^2 \delta \ , \ T(x, y) = (x - x_0) / \exp(y \cdot s) + x_0 \qquad (5.7)$$

where x_0 is the x coordinate of the zero-derivative point of the parabola and δ, s are coefficients. The sign

of s determines the shrinking or expansion of the curve along x direction.

The origin of the coordinate system is at the left mouth corner, and the x,y axes are shown in Fig.5.8. The

coefficients of each parabola curve can be derived from points $\vec{p}_1 \sim \vec{p}_4$ as well as the two mouth corner

points $\vec{p}_l$ and $\vec{p}_r$. The parameters to be optimized are the coordinates of these six control points and the

coefficients for expansion and bending of each curve.

The optimization is performed by maximizing an energy function. The energy function in [Wang et al.

2004] is based on pure region information obtained from clustering. If there are some misclassified pixels

in the probability map, the fitted curve will likely copy this error. We use an energy function as follows

which combines region and gradient information.

$$E = \int_C grad(\vec{p})d\vec{p} + \alpha \int_C region(\vec{p})d\vec{p} \ , \quad region(\vec{p}) = \frac{1}{2} \sum_k \frac{(h(k,\Omega_{ip}) - h(k,\Omega_{op}))^2}{(h(k,\Omega_{ip}) + h(k,\Omega_{op}))^2} \qquad (5.8)$$

where $grad(\vec{p})$ is the normalized gradient magnitude at point $\vec{p}$, $h(k,I)$ is the kth histogram bin value of

image I, Ω_{ip} and Ω_{op} are the regions inside and outside the mouth in a small neighborhood of point $\vec{p}$,

and C includes the entire lip contour. $region(\vec{p})$ defines the region difference in the neighborhood of each

point. α is the coefficient to balance the two terms.
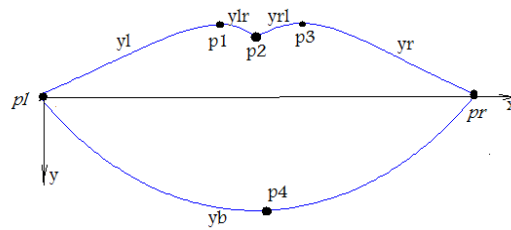
Fig.5.8. Parabola model if the lip: pl, pr: mouth corner points; p1~p4: four points used to determine the parameters of

the parabolas; y*: parabolas for different segments of the lip

5.4.3 Parameter Initialization and Optimization

The rough positions of the control points can be estimated from the initial mouth contour obtained in 5.4.2.

First, the leftmost and rightmost points are identified as mouth corners. After the two corner points are

located, we detect the rough locations of other four control points from the initial contour. The coordinate system in Fig.5.8 is first constructed, with the line connecting the two corner points as the x axis. Under this coordinate system, the middle point in the x direction on the upper lip is first located, and the point in the vicinity of this point with the highest y coordinate is $\vec{p}_2$ (note that y axis is pointing down). $\vec{p}_2$ divides the upper lip into left and right parts. On each part the point with smallest y coordinate is $\vec{p}_1$ and $\vec{p}_3$ respectively. On the lower lip contour the middle point in the x direction is $\vec{p}_4$. Fig.5.7(d) shows the estimated initial positions of the control points.

Other parameters to be optimized are the bending parameter δ and the expansion parameter s for each parabola curve. We set $s = 0$ and $\delta = 0.0001$ for each curve. The optimization is constrained by the following inequalities to ensure reasonable spatial relation among $\vec{p}_1(x_1, y_1)$, $\vec{p}_2(x_2, y_2)$ and $\vec{p}_3(x_3, y_3)$.

$$x_1 < x_2 < x_3, y_2 > y_1, y_2 > y_3 \qquad (5.9)$$

The feature points in Fig.5.2(b) coincide with the control points in Fig.5.8, so the optimized parameters of the parabola curves give the coordinates of the feature points as well. Fig.5.7(e) shows the lip contour and the feature points after optimization.

5.5 Nose Features

Very few literatures have reported feature detection on the nose, probably because the nose does not have distinct feature points as the mouth and the eyes do. In this paper, we are interested in the alare point (al) which is the most lateral point on each side of the nose (Fig.5.2(c)), and the nose boundary passing through the alare point on each side. However, since each boundary is an open contour, there is no clear upper or lower bound of the contour.

5.5.1 Nostril Detection

In order to facilitate the detection of the nose boundary, we first locate the rough positions of the nostrils. The average intensities of the rows and the columns of the image are calculated, generating a horizontal profile and a vertical profile of the intensity (Fig.5.9(a)). The x coordinates of the nostrils can be found at

the local minima on the horizontal profile, and the y coordinate can be found at the local minimum on the vertical profile The search range of the nose boundaries is then limited to the two sides of the nostril points (Fig.5.9(b)).

5.5.2 Nose Boundary Detection

Since the nose boundary is usually weak, histogram equalization is first applied on the grayscale nose image to enhance the contrast (Fig.5.10(a)). An edge map of this enhanced nose image can usually capture most of the nose boundaries. However, it is not always possible to get perfect edges around the nose boundaries, and an example of the missing edge around one boundary is shown in Fig.5.10(b). Alternatively, local minimum map (Fig.5.10(c)) can be applied here as well which is very close to the edge map around the boundary points, but similar problems may occur as with the edge map. Therefore, we decide to fuse the edge map and local minimum map into a combined map that can capture the boundary points more robustly. Specifically, a local minimum point is added to the edge map only if there is no surrounding edge point in a small neighborhood of this local minimum point, because this is likely to be the position where an edge point is missing. Then the edge/local minimum points (or pixels) are grouped by connected components algorithm, and small components are discarded. Fig.5.10(d) shows the combined map from Fig.5.10(b) and (c) after small components removal, where the missing edge is filled in by local minimum points.

Generalized Hough transform for arbitrary shapes is then used to detect the boundary contours from the combined map. A template shape and a reference point need to be defined first, and a lookup table is constructed recording the location of each point on the template shape relative to the reference point. Given the location of the reference point, a shape can be generated from the lookup table, thus each point in the Hough space (i.e., reference point) corresponds to a shape in the image space. Assuming that the rotation and scaling of the shape do not change, we have the coordinates of the reference point as the only parameters. Since these parameters are equivalent to those in the line detection with Hough transform, the algorithm for the curve detection follows the same routine as the line detection. In our implementation, the template for each nose boundary is a generic shape of the nose boundary, and the reference point is chosen as the leftmost (rightmost) point for the left (right) boundary (Fig.5.10(e)). With the results of the nostril locations, the range of the parameters is restricted to the part of the image shown in Fig.5.9(b). Rotation

and scaling parameters can also be included, but the computational cost will be significantly increased. We find that a fixed template shape can serve our purpose as well since the nose boundaries do not vary too much among different individuals, so we do not consider rotation and scaling parameters.

A Bezier curve is then used to fit the detected local minimum points on each boundary. The alare points are then located by finding the left most and rightmost point on each contour respectively (Fig.5.10(f)).
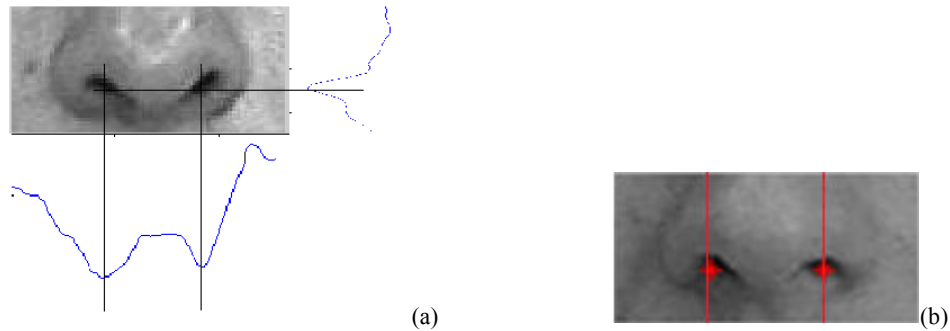
Fig.5.9 (a) illustration of the nostril detection (b) detected nostrils and the search range of the nose boundaries (rectangle regions on the left and right sides)
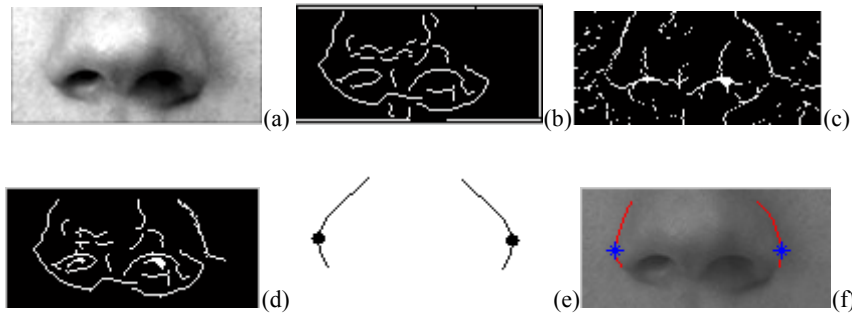
Fig.5.10 (a) nose image after histogram equalization (b) edge map of (a) (c) local minimum map of (a) (d) combined map after small components removal (e) template nose boundaries and the reference points (f) detected nose boundaries and feature points shown on the original nose image

5.6 Chin Features

Previous work on chin contour estimation can be grouped into active contour based methods [Lam and Yan, 1998; Huang and Su, 2002; Sun and Wu, 2002] and parametric model based methods [Kampmann, 2004; Wang and Su, 2003; Chen et al., 2007; Ding and Martinez, 2008]. Parametric models are more robust than active contours because human's prior knowledge about the chin shape is utilized [Chen et al., 2007].

101

Our method falls into parametric model based category. The problem with these methods, however, is the sensitivity to image noise. For example, in [Ding and Martinez, 2008] a quadratic curve is iteratively fit to the edge points, and the edge points far away from the fitted curve are removed in each iteration. This simple curve fitting can be easily misguided by edge noise. [Chen et al., 2007] proposed a modified Canny edge detector to remove most of the unwanted edges, but the result still largely depends on the original Canny edge detection. Our method follows the same idea as for the nose boundary detection, and an additional tensor voting step is applied to enhance salient curves while removing irrelevant noise.

Similar to the nose boundaries, there is no clear upper bound for the chin curve. In our work, we set the top of the mouth as the upper bound for the chin. The feature point of interest is gnathion (gn) which is the lowest point on the curve. The chin localization has to be performed after the mouth feature localization.

5.6.1 Bottom Chin Curve

In this section we detect the part of the chin that is right below the mouth. Similar to [Ding and Martinez, 2008], a bounding box immediately below the mouth bounding box is located (Fig.5.2(d)). The width of this bounding box is the same as the mouth bounding box. A combined map is calculated within this bounding box which is the OR of the local minimum map and the local maximum map along the vertical direction (Fig.5.11 (a-c)). Only vertical direction is considered because this part of chin curve is near horizontal. The reason to use both local minimum and local maximum maps is that the chin points can be intensity valley points or peak points. In this example, the left part of the chin consists of intensity valley points (Fig.5.11 (a)) and the right part consists of intensity peak points (Fig.5.11 (b)). The same as in 5.5.2, small noise regions are removed from the combined map (Fig.5.11 (d)).

Similar to the nose boundary detection, generalized Hough transform is to be applied on the combined map to locate the chin. Although small noise regions have been removed, there are still some discontinued short line segments in the combined map which are part of the chin (Fig.5.11 (d)). Thus, a tensor voting [Tang et al., 2000] step is added prior to applying generalized Hough transform. Tensor voting is a powerful method to identify salient structures in a noisy dataset. Briefly, the input data (points on the combined map in our case) are encoded as elementary tensors. Each tensor communicates with its neighboring tensors by vote casting. Tensors that lie on salient features (e.g., curves in 2D) strongly support each other and deform according to the prevailing orientation. The orientation of each tensor encodes the local orientation of

features, and the size encodes their saliency. Features can be extracted by examining the tensors. The result of tensor voting is shown in Fig.5.11(e), where the salient curves are enhanced and the noise is greatly reduced. The only parameter of tensor voting is the scale, which is set empirically in our work. Although there are some methods about multiscale tensor voting [Loss et al., 2009], we find that one scale can work for our task because we only look for the chin instead of multiple objects.

In the generalized Hough transform, the template curve is defined as a fixed quadratic curve $y = kx^2$ with width w and height h (Fig.5.11(f)). w is the same as the width of the chin bounding box, and the reference point is point (0,0) on this curve. In 5.2.2, the coordinates of the reference point are the only parameters to be solved, but the situation is slightly different here. Due to the near-symmetry of the chin curve along x direction, the x coordinate of the reference point can be fixed at $w/2$. Besides, the height h can vary within a certain range in order to accommodate different chin shapes. Thus, the parameters here are the y coordinate of the reference point and h. As in 5.5.2, a Bezier curve is then fit to the detected chin points (Fig.5.11(g-i)). Since the bottom chin is short and does not vary dramatically among subjects, this simple quadratic curve is sufficient to model its shape. When a larger part of the chin is to be detected, more complex model is needed which is shown in 5.6.2.
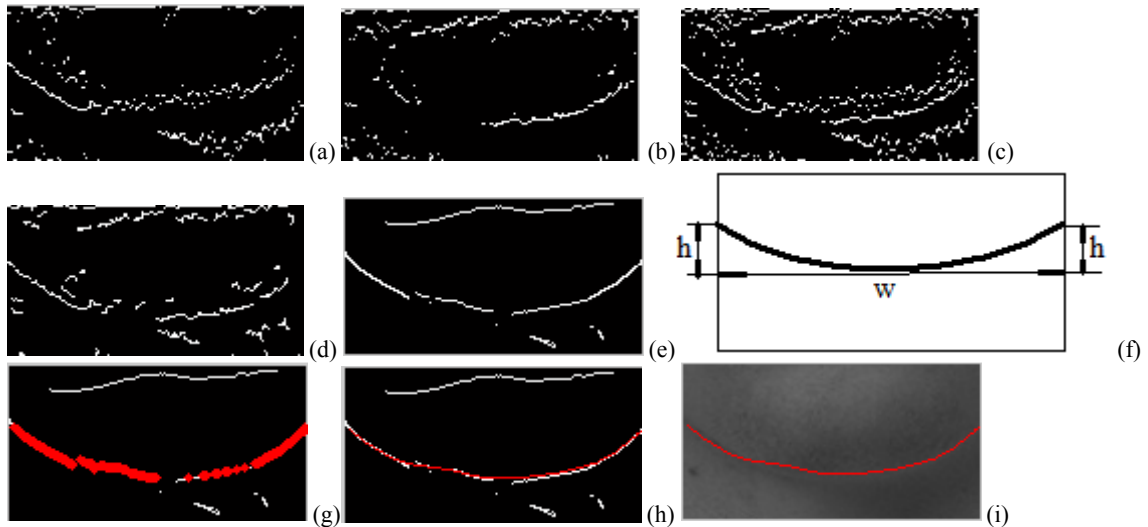
Fig.5.11 (a) local minimum map (b) local maximum map (c) combined map (d) combined map after small noise removal (e) enhanced binary image after tensor voting (f) template curve model (g) detected chin points overlaid on the binary image (h) final chin curve overlaid on the binary image (i) final chin curve overlaid on the original chin image

5.6.2 Extended Chin Curve

Now we extend the chin contour up to the level of the mouth. Note that if the bottom chin is sufficient in a particular application, this section can be skipped. The face image (e.g., Fig.5.3(b)) is cropped from the top of the mouth and below. We divide this cropped image into top and bottom parts, separated at the bottom of the mouth bounding box (Fig.5.12(a)). The bottom image has the same rows as in the image shown in Fig.5.11(i), but more columns on each side. Besides the chin curve detected in 5.6.1, more portion of the chin is visible in this image. The same methods as in 5.6.1 are applied to detect the extended chin curve in the bottom image, except with a more detailed curve model. The template curve consists of two half quadratic curves $y_1 = k_1 x^2$ and $y_2 = k_2 x^2$ connecting at the point (0,0), with height h and widths w1 and w2 (Fig.5.12(b)). The reason for using two quadratic curves is that there is often slight asymmetry between the left and right parts of the chin. The reference point of this template curve is also (0,0). Since the partial chin curve is already obtained in 5.6.1, the location of the reference point can be fixed at the lowest point on the partial chin curve (the red dot in Fig.5.12(a)), and h can be set equal to the distance from this point to the top row of the bottom image. The only parameters in generalized Hough transform are w1 and w2. The binary image after tensor voting and the detected chin points are shown in Fig.5.12(c) and (d).

The chin curve in the top image is relatively simple, consisting of two near straight lines. We apply Sobel vertical edge detector on the top image (Fig.5.12(c)), followed by Hough transform to detect points on the the two lines (Fig.5.12(d)). Finally, a Bezier curve is fit to all the detected points (Fig.5.12(e)).

(a)   (b)

(c)   (d)

104

(e)
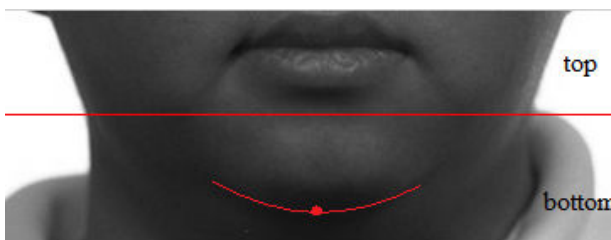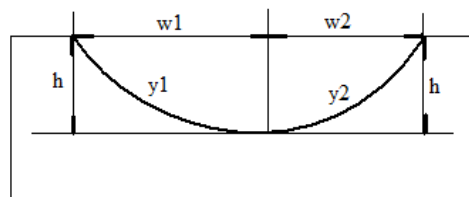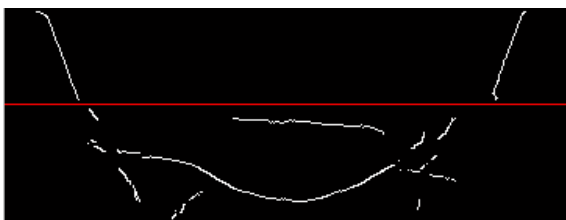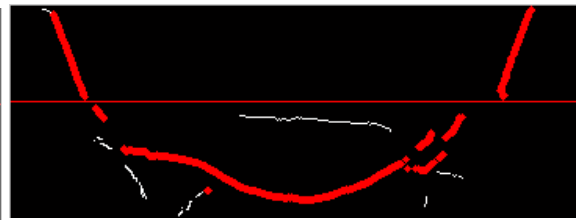
Fig.5.12 (a) top image and bottom image separated by the red line; the detected bottom chin curve and the lowest point marked as a red dot (b) template curve model (c) Sobel edge detection on the top image and tensor voting result on the bottom image (d) detected points overlaid on the image in (c) (e) final chin curve overlaid on the original image

5.7 Results and Validation

The proposed framework is tested on FERET database [Phillips et al., 1998; Phillips et al., 2000], which contains 11338 color images of faces with different expressions, rotations and lighting conditions. From the frontal face subset, the images that obviously violate our assumptions are removed, and 200 images are randomly selected from the remaining images which meet our assumptions to a maximum extent. Slight facial expressions and head rotations are tolerable for our method. Each image has the dimension of 512*768 and includes the full head and the neck of a single person against a near-white background (e.g., Fig.5.3(a)).

Our framework is also tested on BioID database [BioID] which contains 1521 grayscale frontal face images of dimensions 384×286, acquired under various lighting conditions in a complex background. 100 images are selected in the same way as for those in FERET database. Mouth features are not applicable to this database since our mouth feature localization technique requires color images. However, the bounding box of the mouth can still be detected, so we use the top of the bounding box as the upper limit of the chin.

Fig.5.12 shows several results on the detected face images, from which we can see the extracted feature points and boundary contours are faithful to the image data. In order to quantitatively evaluate the performance of our method, the feature points are manually labeled on each image by three different operators. The ground truth position is taken as the average of the manually labeled positions from the three operators. For each feature point on an individual image, the accuracy of our detection result is calculated as

$$a = 1 - \| \vec{p}_g - \vec{p}_t \| / d_w \quad (5.9)$$

105

where $\vec{p}_g$ is the ground truth position of this feature point, $\vec{p}_t$ is the position obtained by our method, $\|.\|$ is the Euclidean distance, and d_w is the width of the corresponding facial component. For example, d_w is the eye width when the accuracy of an eye feature point is calculated, which can be estimated from the distance between the ground truth positions of two eye corners. The value of a is truncated to [0,1]. The result of a feature point is considered successful if its accuracy is above a threshold T, and the result of an entire image is successful if all the feature points are successfully detected. The overall success rate for each feature point is the percentage of the successful cases, and the overall success rate for all feature points is the percentage of successful images.

Among all the results, there is no failure in face detection or region detection. Note that the face detection is considered successful if all the facial components of interest are completely inside the detected face, and the region detection is considered successful if all the related feature points are inside the detected region. Among all the failure cases, the main cause of the mouth failure is the incorrect separation between the lip and the skin due to the lack of color constancy under poor lighting. The failure on the eye features is mostly because the iris is connected to other parts in the binary image thus leading to incorrect iris detection. The failure of the nose features is mainly due to extremely low contrast on that particular boundary. Table 5.1 shows the average accuracy of each feature point, the success rate for each feature point and the overall success rate on FERET dataset. Table 5.2 shows the same measurements on BioID dataset. The success rate of individual feature point remains high even with a high threshold T, although the overall success rate declines faster with the increasing T because it is affected by every single feature point.

No validation of the boundary contours is performed because it is difficult to provide the ground truth by manually delineating the contour. Since most of the feature points are obtained from the contours, the accuracy of the feature point detection can reflect the accuracy of the boundary extraction to some extent.
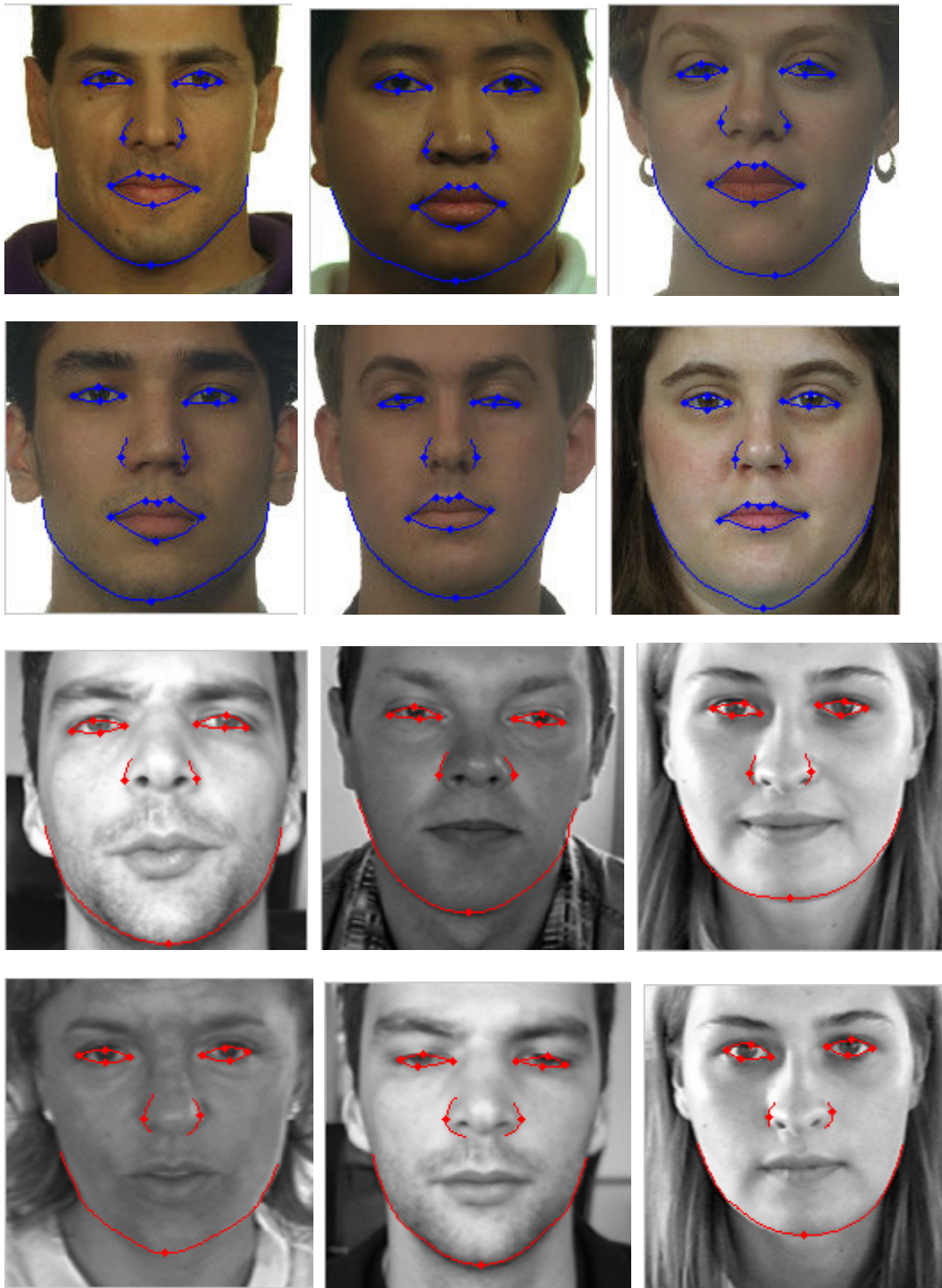
Fig.5.13 Results on FERET database (top two rows) and BioID database (bottom two rows)

Table 5.1 Success rate of each feature point and the overall success rate on FERET database

Facial components	Feature points	Accuracy	Success rate

			T=0.95	T=0.9
Left eye	ex	0.948	0.76	0.97
	en	0.941	0.72	0.95
	ps	0.914	0.71	0.78
	pi	0.936	0.72	0.82
Right eye	en	0.937	0.72	0.96
	ex	0.932	0.78	0.93
	ps	0.910	0.69	0.76
	pi	0.923	0.75	0.89
Nose	al (left)	0.901	0.66	0.76
	al (right)	0.934	0.75	0.78
Mouth	ch (left)	0.975	0.96	0.99
	lss (left)	0.967	0.92	0.98
	ls	0.965	0.90	0.98
	lss (right)	0.962	0.87	0.95
	ch (right)	0.963	0.89	0.96
	li	0.938	0.76	0.94
Chin	gn	0.912	0.70	0.80
Overall			0.60	0.72

Table 5.2 Success rate of each feature point and the overall success rate on BioID database

Facial components	Feature points	Accuracy	Success rate	
			T=0.95	T=0.9
Left eye	ex	0.942	0.74	0.96
	en	0.945	0.78	0.96
	ps	0.908	0.68	0.84
	pi	0.934	0.70	0.86
Right eye	en	0.943	0.74	0.96
	ex	0.938	0.72	0.93
	ps	0.911	0.66	0.83
	pi	0.921	0.68	0.88
Nose	al (left)	0.886	0.62	0.74
	al (right)	0.892	0.65	0.78
Chin	gn	0.916	0.72	0.87
Overall			0.58	0.70

5.8 Conclusion

This chapter proposes a framework for automatic facial feature detection. Features of the eyes, the nose, the mouth and the chin are extracted. The primary application of this work is the computer aided diagnosis, so only frontal images are considered and some other assumptions of the input image are made. Visual inspection of the experimental results shows the accuracy of the overall feature detection, and quantitative validation further demonstrates the accuracy of the feature point detection.

6. Local Shape Analysis of the Eye in Autism Subgroup

As stated in 1.1, ASD is a heterogeneous group of disorders. A homogeneous subgroup called MFP group is hypothesized based on facial gestalt. The purpose of this study is to quantitatively verify the clinical observations of MFP subjects. It is observed that MFP subjects have distinct eye features compared to the rest of ASD population, such as wider inter-eye distance and almond-like eye shape. We are particularly interested in the local shape of the eye since it cannot be measured physically or calculated from a few landmarks, yet it is most observable by visual inspection. In this chapter we will only focus on local shape analysis (4.1.5) of the eye which has not been performed in previous works, although the landmark analysis methods in 4.1.4 can also be applied to the eye shape.

6.1 Subjects and Data

600 children with ASD were evaluated at the Thompson Centre for Autism and Neurodevelopmental Disorders last year. The age of the subjects ranges from 10 to 12. This narrow range is selected so subjects are prepubertal but have completed 90-95% of head growth [Farkas & Posnick, 1992] and brain growth [Dekaban, 1978] and are at the same stage of facial development, which is a continuous process through the seventh decade [Miroue & Rosenberg, 1975]. Children with disorders known to cause an autism phenotype such as Fragile X, chromosome abnormalities or severe prematurity with brain damage and children with IQs/DQs less than 50 will be excluded. Subjects will all be of northern European ancestry. Minority subjects and controls will not be included in this pilot, exploratory study since facial measurements have not been developed for groups other than those of European ancestry [Farkas, 1994]. We anticipate that an important component of future studies will be to study the generalization of the results to all ethnic groups. From this large ASD population, 15 subjects have been identified as MFP group by clinical experts. Only one of the 15 MFP subjects is female. This extreme ratio does not reflect normal male-to-female ratio in ASD population. Therefore, we decide to exclude the female MFP subject in this study to minimize the unpredictable gender effect. Accordingly, 14 males from the rest of the ASD subjects were randomly selected to form the ASD group, and 14 age-matched male control subjects were recruited from the community under the regulations of the Thompson Centre control subject recruitment protocol.

In summary, there are three groups of equal numbered subjects in this study, namely, ASD group that do not show any sign of MFP, a subgroup of ASD that are diagnosed as MFP group, and the control group. In the rest of this chapter, we use 'ASD group' to indicate the non-MFP ASD subjects when the context is clear.

2D facial images of the participants can be obtained by a regular digital camera. In this study, however, the images were collected in conjunction with the acquisition of the 3D face surface intended for future research. We conducted full 360 degree head/face scan using the 3DMD® cranial system (http://3dMD.com) to reconstruct the 3D surface model. Both the geometry and the co-registered texture images were obtained. Only the texture images were used in this study. All images have the size 1234x1624and uniform dark blue background.

6.2 Methods

The method in 5.3 is used to locate the feature points and eyelid contours on the facial images. The images are resized by half in order to speed up the process. To ensure the accuracy of feature locations, all results are examined by a human operator, and manual adjustment is made if necessary. Four landmark points can be obtained which are the two eye corners, the highest point and the lowest point (Fig.5.2(a)).

The local shape analysis described in 4.1.5 is performed here. The eye corners divide the eye contour into two open contours, i.e., upper eyelid and lower eyelid. Although each contour can be further divided by the highest and lowest point respectively, it is unnecessary for the shape comparison because these contours are already simple enough. A local coordinate system is built as shown in Fig.6.1, and it can be used for both upper eyelid and lower eyelid since they share the end points. The distance between eye corners is also normalized to one. Similar to 4.1.5, each eyelid contour is uniformly sampled with equal number of points on all images. The local shape comparison of both eyelid contours can be conducted at the same time since both contours share the same coordinate system. Three groups of comparison are performed, i.e., between MFP and ASD, MPF and control, control and ASD. The statistical method is the same as in 4.1.5. Only the subject's right eye is compared in this study.
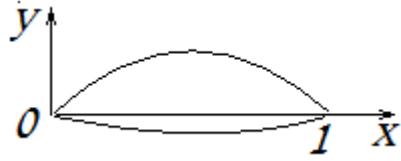
Fig.6.1 Local coordinate system of the eye

6.3 Results

Fig.6.2(a)-(c) each shows the average eye shapes in one comparison, and Fig.6.2(d) shows the average shapes of the three groups overlaid together. From visual inspection of the average shapes, we find that the average shapes of the controls and MFP do not differ very much (Fig.6.2(c)), while there is larger difference between the shapes of ASD and controls (Fig.6.2(a)), and between ASD and MFP (Fig.6.2(b)).

Fig.6.3 shows the significance maps in each comparison. A significance level of 0.05 is selected. Fig.6.3(a) shows that ASD and controls differ significantly in the upper and lower eyelids around the outer eye corner, Fig.6.3(b) shows that ASD and MFP differ significantly in the middle upper eyelid as well as the lower eyelid around the outer corner, and Fig.6.3(c) shows almost no significant difference between the eye shapes of MFP and controls.
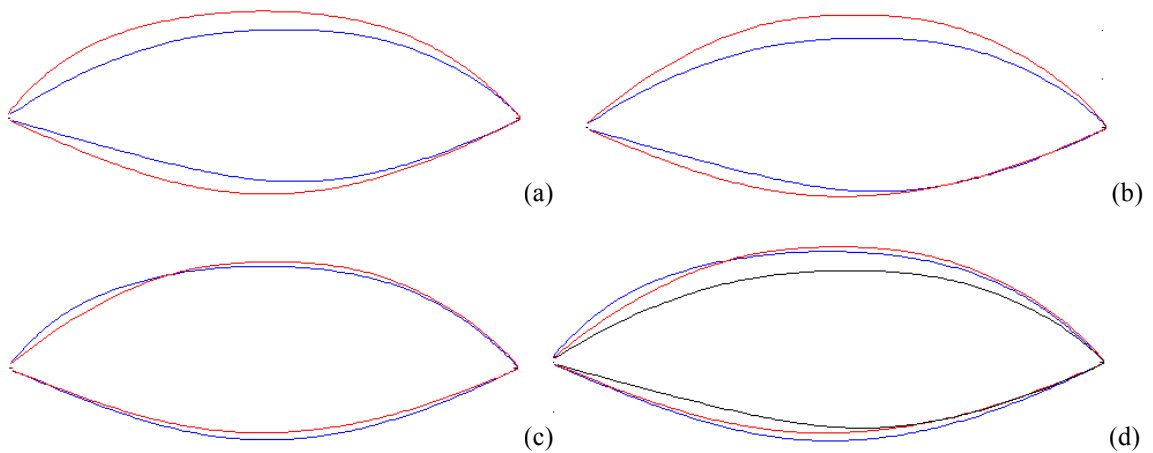
Fig.6.2 Average eye shapes of the groups (a) ASD (blue) and control (red) (b) ASD (blue) and MFP (red) (c) control (blue) and MFP (red) (d) ASD (black), control (blue) and MFP (red)

112
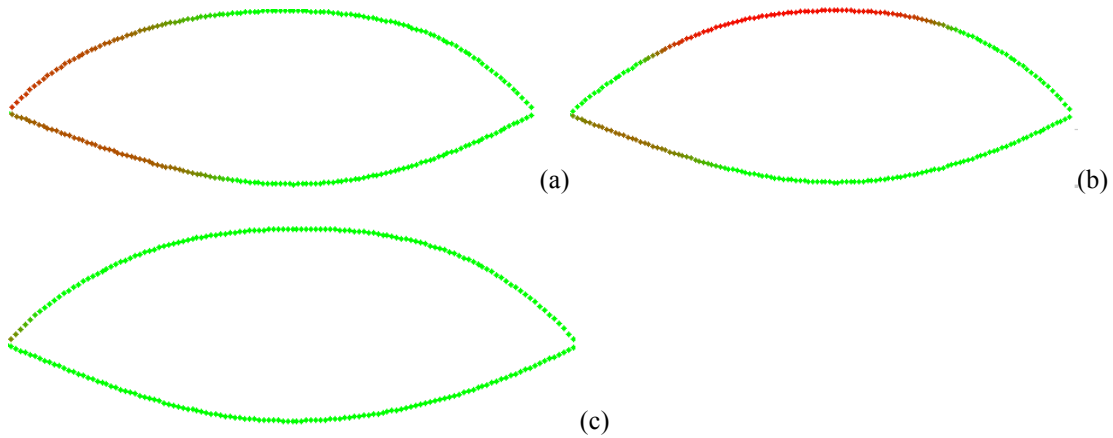
Fig.6.3 Significance maps of the three comparisons (a) ASD and control (b) ASD and MFP (c) control and MFP

6.4 Conclusion

This study aims at verifying the hypothesis about the facial gestalt of a subgroup in ASD population named MFP group. The clinical observation states that MFP subjects have distinct facial phenotypes, especially in the eyes. This conclusion is drawn from the observation within ASD population. However, it is unclear whether there is any difference in the facial phenotype between ASD and normal subjects. Therefore, this study includes control subjects in the comparison as well as the ASD (non-MFP) and MFP subjects. Although it is a pilot study on the right eye shapes of a small number of subjects, the findings of this study are very interesting. As expected, there is significant eye shape difference between ASD and MFP. In addition, we find significant difference between ASD and controls. The most interesting finding is that there is no significant difference between MFP and controls. Fig.6.2(c) shows that the average shapes of these two groups almost overlap. This indicates that the eye shape of the MFP group is actually more 'normal' compared to that of the rest of ASD population. The observed eye shape difference between MFP and the rest of ASD population is due to the distinct eye shape in the majority of ASD subjects (i.e., non-MFP). Fig.6.2 shows that the eye shapes of both control group and MFP group are more rounded than those of ASD group, especially the upper eyelid. This explains the clinical observation that the eyes of MFP group look more 'almond-like' compared to the rest of ASD. However, this feature is not exclusive for MFP, because normal controls also have a similar eye shape.

Based on this study, we can draw two conclusions. First, the majority of ASD subjects have a different eye shape than the controls. The difference is that the eyes of the controls are more rounded, and the eyes of

ASD subjects are thinner around the outer corner. Second, a small portion of ASD population do not have this distinct eye shape as the rest of the ASD subjects, so their eye shapes look similar to the controls. This portion of ASD population is called MFP group, which is a group that looks more 'normal' than the rest of the ASD population. Therefore, MFP group may have a milder facial phenotype than other ASD subjects, although further study on other facial components needs to be performed to confirm it.

The results of this study need to be interpreted with caution due to several limitations. The sample size is relatively small, which may affect the generality of this study. Only male subjects are considered in this study, so the conclusion may not apply when female subjects are included. However, sexual dimorphism is ruled out in this way, which can be an advantage of this study. Despite these limitations, this study still provides valuable insight on the facial phenotype of MFP group. Future work on a larger sample size and other facial components remains to be done.

7 Conclusion and Future Work

7.1 Work Summary

The development of computational technologies has potential benefits in medical field, such as reducing labor force, increasing computational accuracy and performing complex tasks that are impossible for humans. However, the real application of computational technologies in medical field is still facing great challenges. It is a long-term goal and this thesis demonstrates some efforts towards this goal. Particularly, it focuses on the analysis of the brain and the face in the context of ASD population, although the methods can be applied in many other medical contexts.

One interest from the neurological community is how the brains of the neurodevelopmental disordered people differ from normal, which is the aim of brain morphology analysis. The differences can be in size or shape, while shape differences are more subtle and challenging. In this thesis, we focus on the shape morphology of a few brain structures that are of interest in our clinical context, such as the corpus callosum. Since individual brain structures need to be compared among different subjects, brain structure segmentation serves as a necessary preprocessing step. In Chapter 2, we propose several automatic/semiautomatic segmentation methods for brain structures, which constitute our segmentation framework. These methods are all based on deformable models, but they differ in the model initialization so that each method is more efficient for particular applications. In section 2.1, a context-sensitive active contour method is proposed for 2D corpus callosum segmentation. Prior knowledge of the corpus callosom shape is incorporated in the initial contour. In section 2.2, a semiautomatic 3D segmentation method is proposed which combines region based and boundary based methods in a sequential order. Region based methods are used to initialize the model followed by boundary based model deformation. Since region based methods can generate a close-to-target initial model and are much faster, the computational time in the subsequent model deformation is greatly reduced. Section 2.3 proposes a fully automatic 2D segmentation framework which shares the same idea as in 2.2, but the seed initialization is fully automated by some hierarchical inference scheme. We also developed a comprehensive piece of software for 3D segmentation described in section 2.2, and detailed descriptions are provided in Appendix.

Brain morphology analysis includes the global traits analysis and shape analysis, and we focus on the methodology of shape analysis. We divide shape analysis into three types: landmark analysis, pseudo-landmark analysis and whole surface analysis. Landmark analysis is relatively simple since only a few anatomical landmarks are considered. When the entire shape contour/surface is considered, however, point correspondence must be established before shape comparison. In Chapter 3, we propose two methods for 3D shape correspondence, which is essential in pseudo-landmark analysis and whole surface analysis.

The entire framework for brain morphology analysis is shown in Chapter 4 through several clinical studies, where the methods in both Chapter 2 and Chapter 3 are applied. Since global traits analysis and landmark analysis are relatively straightforward, they are only described in the related applications in Chapter 4 without detailed explanation. The major clinical application is the brain morphology of ASD children, while studies on other populations are also included. Although significant shape differences between comparison groups have been found in our studies, there are some limitations in these studies especially the small sample size. Therefore, the results in these studies need to be interpreted with caution, and further studies on more data are desired in the future.

Another important clue of neurological disorders is the facial patterns, which had not drawn much attention until recently. Unlike brain MRI, facial images are much easier to obtain and less expensive. Therefore, computer based diagnosis based on facial images has gained more and more interest from clinicians. The most challenging technique is the automatic localization of facial features from images. In Chapter 5, we propose a set of methods to locate facial features on the eyes, the nose, the mouth and the chin. These methods form a framework for automatic facial feature localization. Not only feature points, but also feature contours are located, which makes point-by-point shape comparison of facial components feasible. In Chapter 6, we conduct the eye shape analysis among ASD, MFP and control groups. The purpose is to verify the hypothesis about the distinct facial phenotype in MFP group which is a subgroup of ASD. The methods in Chapter 5 are applied to delineate the eye shape contour. The shape analysis methods in Chapter 4 can be directly translated to the eye shape analysis in Chapter 6. Again, small sample size is a limitation of the study in Chapter 6, although some interesting results have been found.

7.2 Potential Value and Future Work

As stated in Chapter 1, the goal of our work is to bridge the gap between computational methodology and clinical applications. Besides the clinical applications presented in this thesis, our computational framework has its potential value in many other medical fields, which gives possible future directions of our work. The methods in Chapter 4 can be translated to the shape analysis of other anatomical structures in other populations of interest. Note that the particular methods for image segmentation may need to be changed according to the target structure, image modality and other application-dependent factors, while the shape analysis framework (Fig.4.15) can be directly applied. Another promising field, though less obviously related, is the treatment planning in radiation therapy. Usually, three dimensional images of the patient's anatomy are taken using medical imaging technologies such as computer tomography (CT). Image segmentation methods are then applied to highlight structures of interest, such as the tumor and the critical structures surrounding it. Dose distribution can then be optimized based on the spatial relation between the tumor and the critical structures, as well as the size and shape of the tumor. Therefore, when incorporated with more domain knowledge of radiation therapy, our methodology of image segmentation and shape analysis will have potential value in assisting treatment planning.

Our facial morphology framework will likely see its future in other biometric related diagnosis. Biometric features have been widely used in people identification, and the details are beyond the scope of this thesis. However, certain diseases can also be reflected through biometric features, such as the face, the hand, the ear, and the tongue. Chapter 6 shows a pilot study on the eye shape, and similar approaches can be applied to other biometric features in other diseases, with the modification of feature localization methods. Moreover, automatic diagnosis based on biometric features is also a promising direction to extend our work. Unlike brain MRI or other medical image modalities, most biometric images are easy to acquire under a fast and low-cost setting, which makes computer based diagnosis more practical. Our shape analysis framework paved the way for computer based diagnosis since it proved the clinical hypothesis about the phenotype difference. In the future, machine learning techniques can be incorporated into this framework to achieve automatic diagnosis.

Appendix A: User Guide on Semiautomatic 3D Segmentation Software

This appendix provides a user guide on how to use the 3D segmentation software described in 2.2. The software, named Brain Explorer, is developed by Qing He and Kevin Karsch.

A.1 Overview of the Software GUI

There are six parts of the GUI (Fig.A.1) as follows.

Part 1: image and model visualization

Three orthogonal views of the MRI data as well as a 3D view are provided. The user can select the slice to view using the scroll bar below each 2D view. The user can also rotate the 3D view by scrolling the three rotation bars at the bottom. The view selection panel to the right of the 3D view is used to select the active view. The user needs to specify a view as the active before doing any interaction on that view (e.g., drawing regions, clicking a point).

Part 2: seed initialization tools

These components are designed for region based initialization. The user can set parameters and perform step-by-step operation.

Part 3: deformation tools

These components are mainly used for parameter setting of the model deformation.

Part 4: manual adjustment tools

This part provides several options for manual adjusting the segmentation results.

Part 5: mesh extraction

This part contains a mesh extraction button and toggle buttons to switch between model display and image display.

Part 6: menus

This part contains several menus. Under 'File' menu there are several options to save the result. The 'View' menu is used for zoom-in and zoom-out of the 3D view. The 'Compute' menu is designed for computing statistics of the segmentation result, which is reserved for future use.
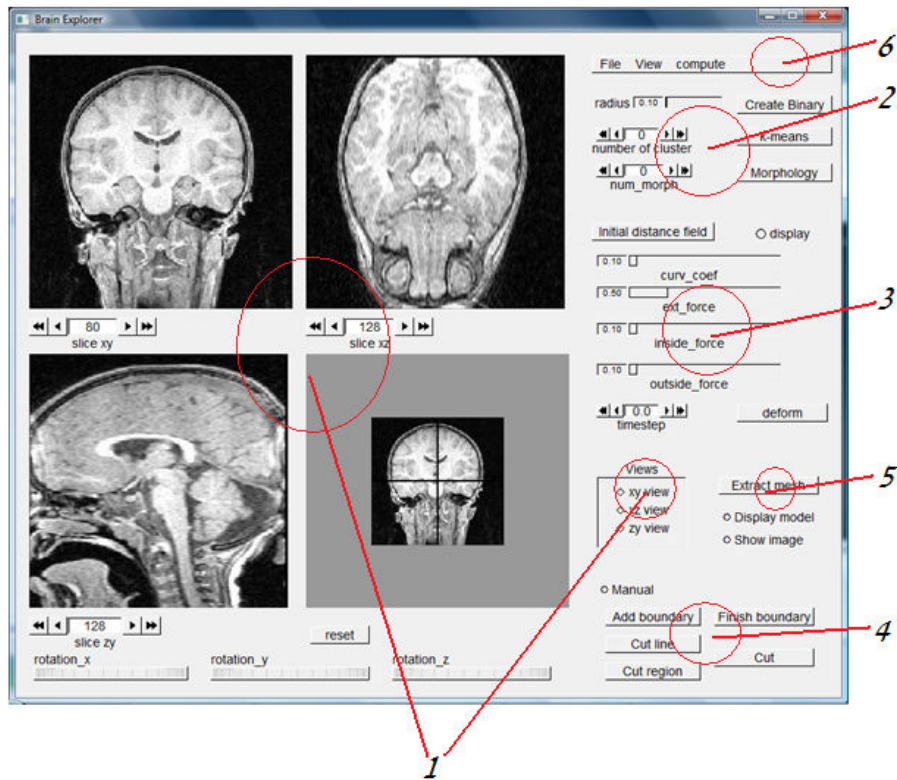
Fig.A.1 Software GUI and its six parts

A.2 Specify the Data

The accepted input data format is Analyze. Other types of data can be converted to Analyze format using software such as MRICro. The file name of the input data can be changed in main.cpp file in the source code. Fig.A.2 shows part of the code with a circle highlighting the file name.

```cpp
int main(int argc, char** argv)
{

    image = new AnalyzeImage();

    image->importAnalyze("data\\sub023");
    cout<<"File loaded"<<endl;

    vminx=0; vmaxx=image->dimX; vminy=0; vmaxy=image->dimY; vminz=0; vmaxz=image->dimZ;
    vminx1=0; vmaxx1=image->dimX; vminy1=0; vmaxy1=image->dimZ; vminz1=0; vmaxz1=image->dimY;
    vminx2=0; vmaxx2=image->dimZ; vminy2=0; vmaxy2=image->dimY; vminz2=0; vmaxz2=image->dimX;
```

Fig.A.2 Specifying the input data in the source code

A.3 Seed Generation

119

The first thing the user needs to do is define a 3D ROI that contains the target structure. This can be done by drawing a rectangle region in each of the two orthogonal views. Any two views can be used to define the ROI. Note that an active view needs to be selected on the view panel before drawing the rectangle. Fig.A.3(a) shows the ROI for the ventricle. The user defined ROI will also be shown in the 3D view. After that, the user needs to right click a point inside the target structure. If the user clicked point is within the ROI, it will show 'seeded 1' in the command window (Fig.A.3(b)). If the user accidentally clicks a point outside the ROI, nothing will show in the command window, which tells that the clicked point is invalid and the user needs to click again.

This program allows two types of user initialization—point initialization and circle initialization. For point initialization, the user can skip the first row of components in part 2 ('radius' sliding bar and 'create binary' button). After setting the number of clusters and pushing 'k-means' button, k-means clustering is performed and the selected cluster will be shown in blue in each orthogonal view. Then morphology can be performed in a similar way. Fig.A.3(c) shows the result after morphology. For circle initialization, the user needs to set the radius of the circle using 'radius' sliding bar. This circle will be centered at the user clicked point. Then pressing 'create binary' button will generate the binary image. After this, the user can perform l-means and morphology operations in the same way as in point initialization.

After morphology, the user needs to push 'initial distance field' button to calculate the distance field. The toggle button 'display' switches between displaying the distance field (Fig.A.3(d)) or the cluster.

A.4 Model Deformation

The four sliding bars in part 3 are used for adjusting the parameters in the deformation equation. The scroll bar below them is for setting the time step. Each time the user presses 'deform' button, 20 iterations of deformation are performed. The user can continue to press 'deform' button to allow more deformation. The model during deformation is shown in blue (Fig.A.4).

(a)

(b)

(c)

(d)
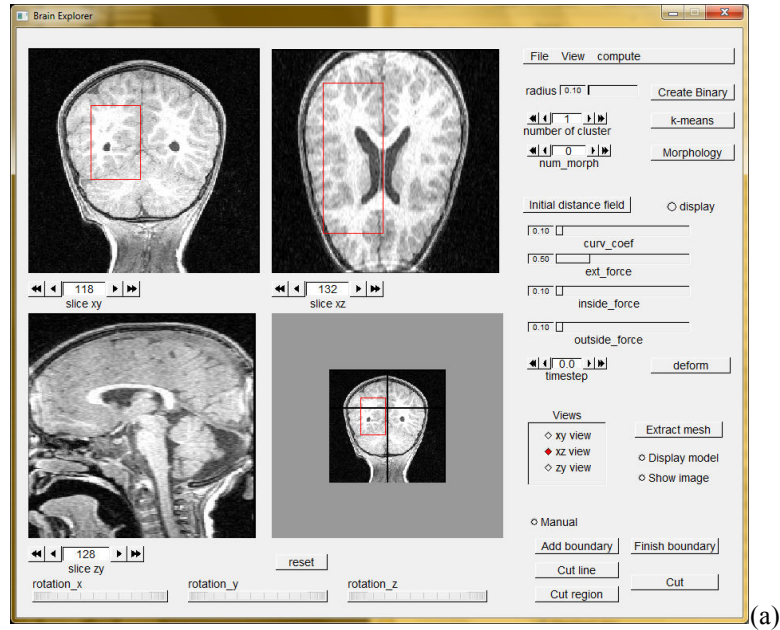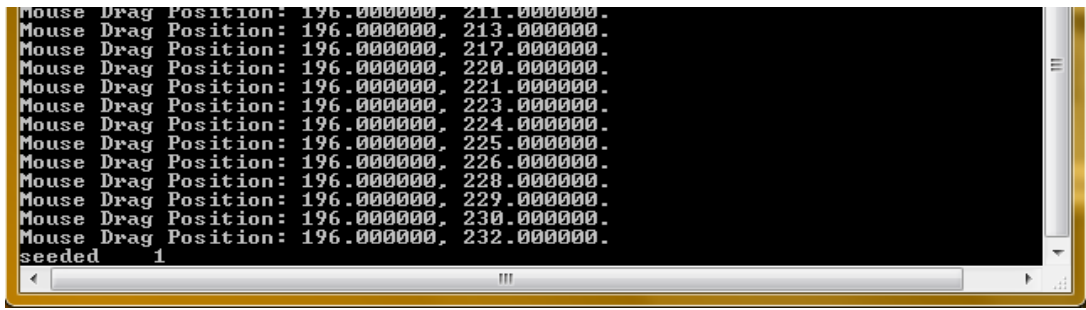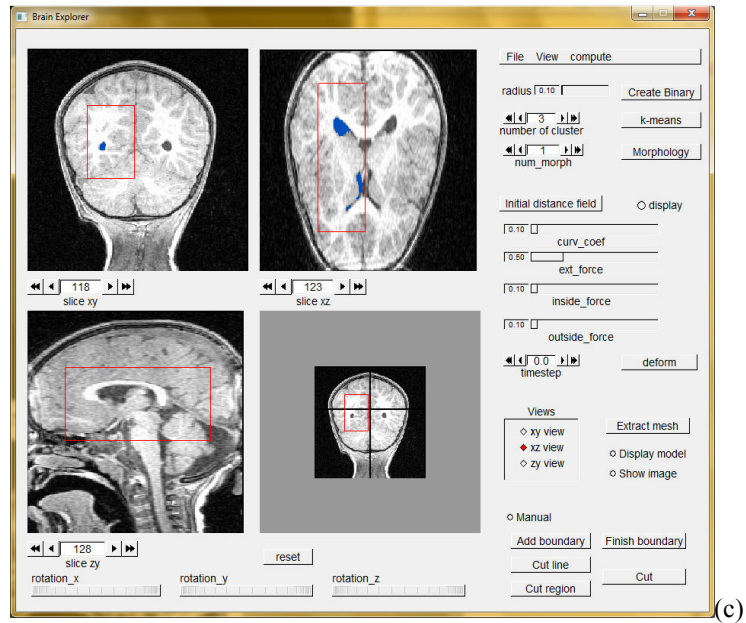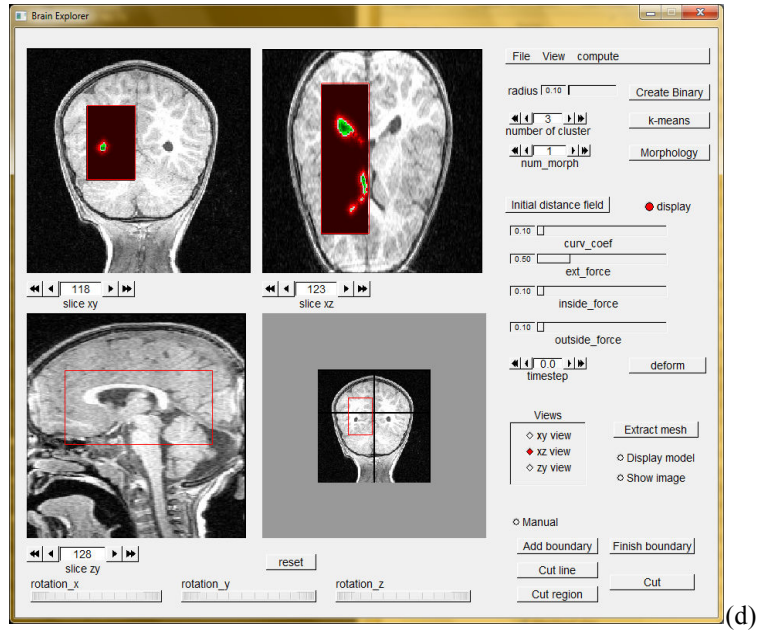
Fig.A.3 (a) ROI of the left ventricle (b) command line window after a correct seed point is clicked by the user (c) the result (in blue) after morphology (d) the initial distance field
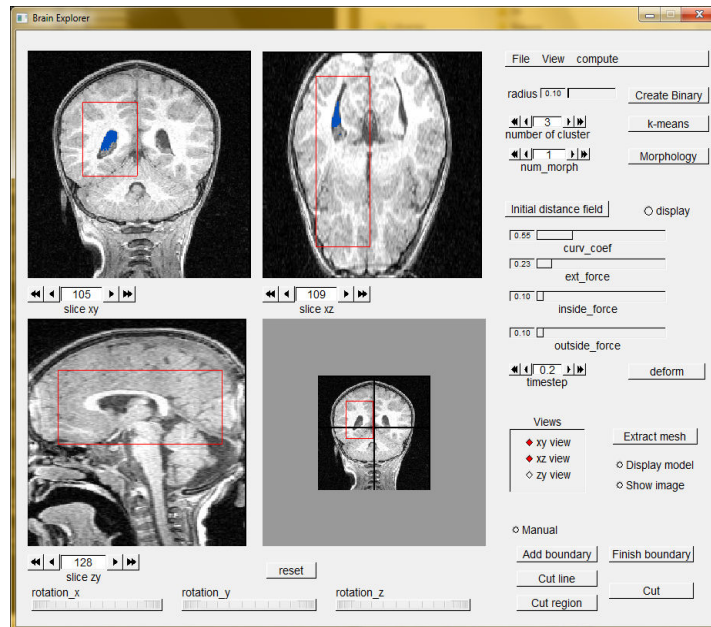
Fig.A.4 The result after deformation

A.5 Manual Adjustment (optional)

This step is needed only when the segmentation result has noticeable defect. Two types of manual adjustment are allowed—fill in the missing part and delete the unwanted part. By selecting the toggle

button 'manual', the user switches the software into manual adjusting mode. The manual adjustment is on a slice by slice basis, so an active view needs to be selected before any manual interaction. If a missing part needs to be added, the user can perform the following steps.

1) click 'add boundary' button after selecting the active view

2) sequentially click a set of points on the missing boundary, in a counterclockwise fashion

3) click 'finish boundary' button, and the missing portion enclosed in the user defined boundary will be displayed in blue.

If an unwanted part needs to be removed, the user can choose to cut off this part by a line or a region. The following steps are for line cutting.

1) click 'cut line' button.

2) click two points that form a line separating the valid portion and the unwanted portion. A red line will be shown on the view (Fig.2.14(c)).

3) click a point on one side of the line indicating which side the unwanted portion lies in. The exact position of this point does not matter.

4) click 'cut' button, and the unwanted part on one side of the line will be cut off (Fig.2.14(d)).

 The following steps are for region cutting.

1) click 'cut region' button.

2) draw a rectangle that encloses the unwanted part by dragging the cursor. A red rectangle will be shown on the view (Fig.2.14(e)).

3) click 'cut' button, and the unwanted part inside the rectangle will be cut off (Fig.2.14(f)).

A.6 Mesh Extraction

The 'extract mesh' button is used to compute the triangle mesh from the implicit representation of the segmentation result. This will take some time to complete. After this, the 3D model can be displayed in the 3D view together with the image. The toggle buttons below 'extract mesh' button are used to select which should be displayed (the model or the images or both). Fig.A.5 shows the resulting mesh in the 3D view

A.7 Save/Load Results

The result can be saved in different formats under 'File' menu. Fig.A.5 shows the drag-down of 'File' menu. There are three menu items for saving the results. 'save obj' saves the 3D mesh in obj format. 'save dat' saves only the 3D points without triangle mesh information. 'save implicit' saves the result as binary volume data. In addition, the 'load' item allows the user to load a previously saved result, and the 'save ROI' item saves the user defined ROI.
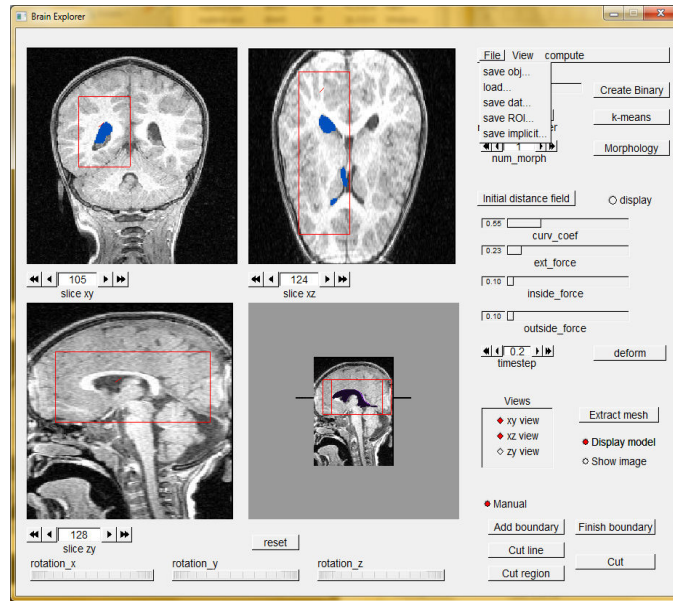
Fig. A.5 The extracted mesh in the 3D view and the drag-down of the 'File' menu

Appendix B: Source Code Structure of the Semiautomatic 3D Segmentation Software

This appendix provides the coding structure of Brain Explore software for developers' guide. The source code is written in VC++ 6.0, and the GUI is built using FLTK.

B.1 File Description

The workspace file is named as dynamic-flow.dsw. Fig.B.1(a) shows the File view of this workspace under VC++ 6.0 environment.

1) Main file

File name: main.cpp

Description: the main( ) function is defined in this file, which is entrance of the program. The file name of the image data and some initial settings are specified in main( ) function.

2) Image analysis files

Files names: analyzeimage.h, analyzeimage.cpp

Description: all image segmentation methods are defined in these files, including clustering, deformation and manual post-processing.

3) Mesh extraction files

File name: MarchingCubes.h, MarchingCubes.cpp

Description: Marching Cube algorithm is defined in these files. These methods are only called in mesh extraction step after the segmentation.

4) Visualization files

File names: view.h, view1.h, view2.h, view3d.h, view.cpp, view1.cpp, view2.cpp, view3d.cpp

Description: files view.*, view1.*, and view2.* correspond to the xy view, xz view and zy view respectively. View3d.* specifies the 3D view. Each .cpp file contains the methods for displaying images/models/user interactions using OpenGL.

5) GUI files

File names: newui.h, newui.cxx

Description: these files are automatically generated by the FLUID application, which is an UI builder of

FLTK. All GUI components and callbacks are defined in these files.

(6) Other files

External Dependencies contain the header files of the GUI components generated by FLTK.
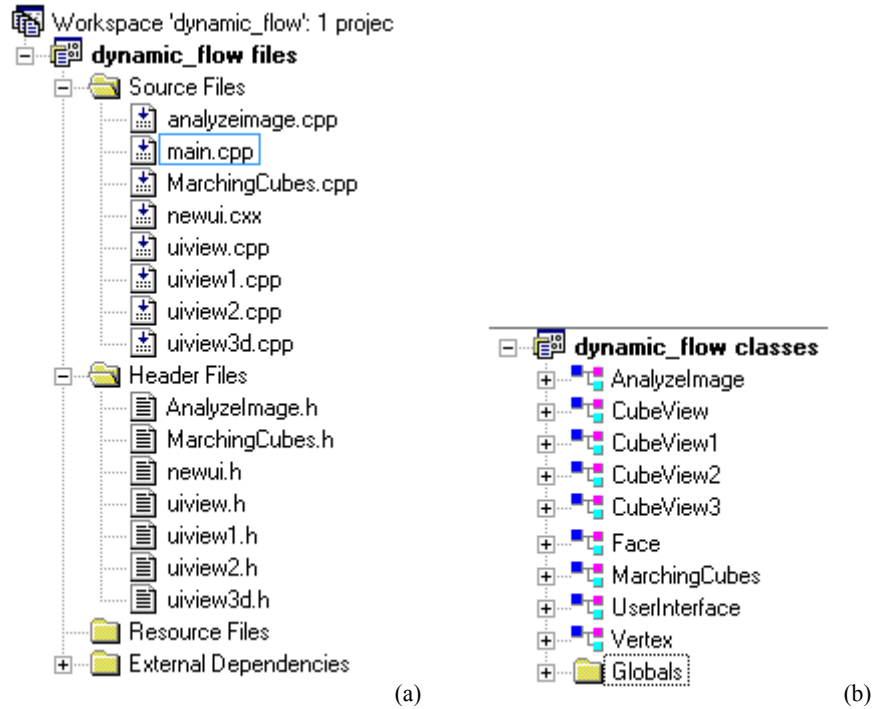
Fig.B.1(a) File view of the workspace dynamic-flow.dsw (b) Class view of the workspace dynamic-flow.dsw

B.2 Class Description

Fig.B.1(b) shows the Class view of the workspace. Similar to the files, these classes can be grouped as

follows.

(1) Image analysis class

Class name: *AnalyzeImage*

Description: this is the most important class in the workspace. Most image segmentation related methods

are defined in this class. The structure of this class is show in Fig.B.2.

```
                                    ┌─arrays: image data, cluster, distance field, etc
                 member variables ──┤
                                    └─scalars: ROI position, seed location, image dimensions, binary indicators, etc
AnalyzeImage ──┤
                                    ┌─ constructor: load image data, initialize member variables
                                    ├─utility functions: sort, search, etc
                 member functions ─┼─ clustering functions: K-means, morphology, connected component, etc
                                    ├─deformation functions: initialize/deform distance field, mesh extraction, etc
                                    ├─manual processing functions: add region, cut region, etc
                                    └─I/O functions: load/save results
```

Fig.B.2 Structure of *AnalyzeImage* class

(2) Mesh extraction class

Class name: *MarchingCubes*

Description: this class contains methods to implement Marching Cube algorithm. An instance of this class

is created in one of *AnalyzeImage* member functions to compute the triangle mesh.

(3) Visualization classes

Class name: *CubeView, CubeView1, CubeView2, CubeView3*

Description: each class corresponds to xy view, xz view, yz view and 3D view respectively.

(4) GUI class

Class name: *UserInterFace*

Description: this class contains a set of UI components and methods generated by FLTK.

(5) Data structures

Structure name: *Face, Vertex*

Description: these structures are used in both *MarchingCubes* and *AnalyzeImage* classes. *Face* defines a

triangle face on the segmented 3D model, and *Vertex* defines a 3D vertex on the model.

(6) Globals

Main() function and other global functions and variables.

B.3 Building the GUI

The GUI of this software is developed using FLTK. FLTK is a cross-platform C++ GUI toolkit for

UNIX[®]/Linux[®] (X11), Microsoft[®] Windows[®], and MacOS[®] X. FLTK provides modern GUI functionality

without the bloat and supports 3D graphics via OpenGL[®] and its built-in GLUT emulation.

FLTK also includes an excellent UI builder called FLUID that can be used to create applications in minutes. Fig.B.3(a) shows the interface of FLUID that displays all the UI components in this software, and Fig.B.3(b) shows the layout of the GUI at development stage. To create a new GUI or load a previously saved GUI, choose File->New or File->Open respectively. To add a component on the GUI, choose a UI type under New menu in Fig.B.3(a). File->Write code will automatically generate the code of GUI components in associated workspace files.

Fig.B.3 (a) FLIUD interface (b) GUI layout at development stage

References

Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E.. 1992a. Fiber composition of the human corpus callosum. Brain Research 598, 143–153.
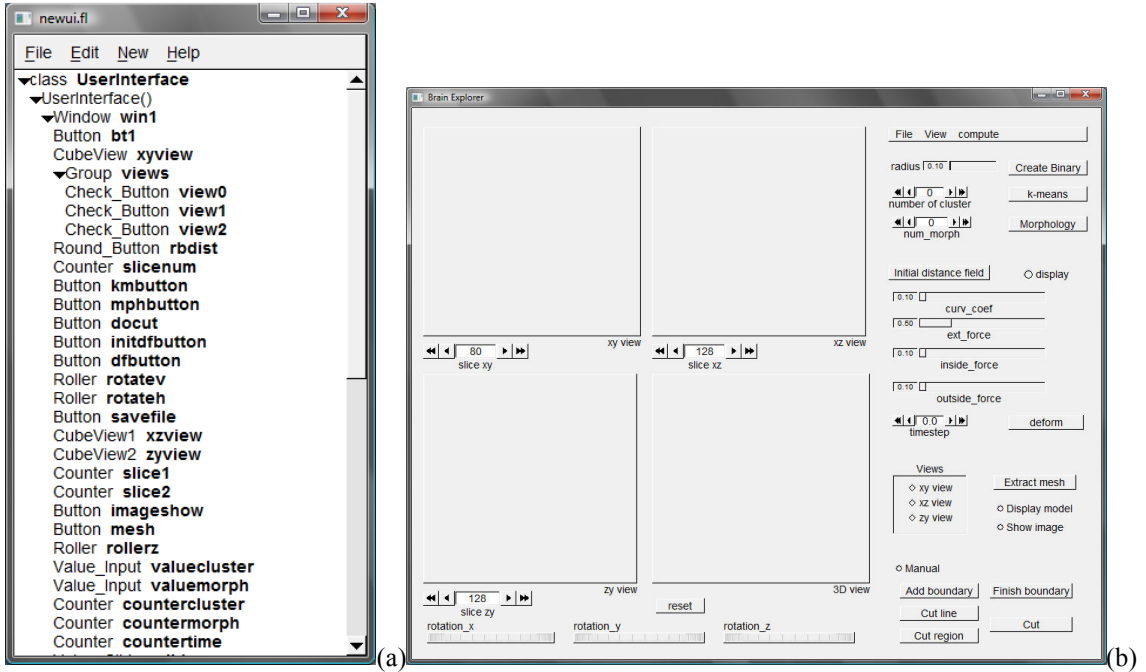
Aguirre, M., Linguraru, M., Marias, K., Ayache, N., Nolte, L.P., Ballester, M.. 2007. Statistical shape analysis via principal factor analysis. In: IEEE Int. Symp. Biomedical Imaging

American Psychiatric Association. 2000. Diagnostic and Statistical Manual of Mental Disorders, 4th ed., text revision, American Psychiatric Association, Washington DC

Asteriadis, S., Nikolaidis, N., Pitas, I., 2009. Facial feature detection using distance vector fields. Pattern Recognition 42, 1388 – 1398

Banich MT, Passarotti AM, White DA, Nortz MJ, Steiner RD. 2000. Interhemispheric interaction during childhood: II. Children with early-treated phenylketonuria. Dev Neuropsychol, 18(1):53-71.

Bello M., Ju T., Carson J., Warren J., Chiu W., A.Kakadiaris I.. 2007. Learning-based segmentation framework for tissue images containing gene expression data. IEEE Transactions on Medical Imaging, 26(5)

Belyaev A. G., Anoshkina E. V., Yoshizawa S., Yano M., 1999. Polygonal curve evolutions for planar shape modeling and analysis. International Journal of Shape Modeling, 5(2): 195-217

BioID: http://www.bioid.com/downloads/facedb/facedatabase.html.

Bhuiyan M.A., Ampornaramveth V., Muto S., Ueno H., 2003. Face detection and facial feature localization for human–machine interface, National Institute of Informatics Journal 5 (3): 25–39.

Boehringer, S., Vollmar, T., Tasse, C., Wurtz, RP., Gillessen-Kaesbach, G., Horsthemke, B., Wieczorek, D., 2006. Syndrome identification based on 2D analysis software. Eur J Hum Genet. 14(10), 1082-9.

Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transaction on Pattern Analysis and Machine Intelligence 11, 567–585.

Bookstein, F.L, 1991. Morphometric tools for landmark data. Cambridge University Press, Cambridge.

Bookstein, F.L., Sampson, P.D., Streissguth, A.P., Connor, P.D., 2001. Geometric morphometrics of corpus callosum and subcortical structures fetal-alcohol-affected brain. Teratology 64, 4–32.

Brigger P., Unser M.. Multi-scale B-spline snakes for general contour detection. 1998. Proceedings of SPIE: Wavelet Applications in Signal and Image Processing VI, 3458: 92-102

Bueno G, Musse O, Heitz F, Armspach JP. 2001. Three-dimensional segmentation of anatomical structures in MR images on large data bases. Magnetic Resonance Imaging, 19(1):73-88

Burrows, A.M., Richtsmeier, J.T., Mooney, M.P., Smith, T.D., Losken, H.W., Siegel, M.I., 1999. Three-dimensional analysis of craniofacial form in a familial rabbit model of nonsyndromic coronal synostosis using Euclidean distance matrix analysis. The Cleft Palate-Craniofacial Journal 36, 196–206. 0:825-830.

Canzlerm, U., Dziurzyk, T.. 2002. Extraction of Non Manual Features for Video based Sign Language Recognition. In: Proceedings of IAPR Workshop, pp. 318-321

Caselles V., Catte F., Dibos F., 1993. A geometric model for active contours in image processing. Numerische Mathematik. 66, 1-31

Caselles V., Kimmel R., Sapiro G., 1997. Geodesic active contours. Int. J. Comput. Vis. 22(1), 61-79

Cates J, Fletcher P, Whitaker R. Entropy-Based Particle Systems for Shape Correspondence. 2006. Int Conf Med Image Comput Comput Assist Interv. 9(WS):90-99.

Cates J, Fletcher P, Styner M, Shenton M, Whitaker R. 2007. Shape modeling and analysis with entropy-based particle systems. Inf Process Med Imaging, 20:333-45.

Chakraborty A., Duncan J.S.. 1995. Integration of boundary finding and region-based segmentation using game theory. In Y. Bizais et al., editor, Information Processingin Medical Imaging, pages 189–201. Kluwer

Chan T.F., Vese L.A.. 2001. Active Contours without Edges. IEEE Transactions on Image Processing, vol.10, no.2

Chen Q., Cham W., Lee K.. 2007. Extracting eyebrow contour and chin contour for face recognition. Pattern Recognition. 40(8): 2292-2300, Part Special Issue on Visual Information Processing

Chen T, Metaxas D. 2005. A hybrid framework for 3D medical image segmentation. Medical Image Analysis, 9(6):547-65

Chou, Y., Lepore, N., de Zubicaray, G. I., Rose, S. E., Carmichael, O. T., Becker, J. T., Toga, A. W., Thompson, P. M., 2007. AUTOMATED 3D MAPPING & SHAPE ANALYSIS OF THE LATERAL VENTRICLES VIA FLUID REGISTRATION OF MULTIPLE SURFACE-BASED ATLASES, Biomedical Imaging: From Nano to Macro, 4th IEEE International Symposium, 1288-1291

Csernansky, J. G., Wang, L., Jones, D., Rastogi-Cruzk, D., Posener, J. A., Heydebrand G., Miller, J. P., Miller, M. I., 2002. Hippocampal Deformities in Schizophrenia Characterized by High Dimensional Brain Mapping, Am J Psychiatry, 159, 2000-2006

Clarke, J.M., McCann, C.M., Zaidel, E., 1998. The corpus callosum and language: anatomical--behavioral relationships. In: Right hemisphere language comprehension: perspectives from cognitive neuroscience, Beeman, M., Chiarello, C. (Eds.), Lawrence Erlbaum, Mahwah, pp. 27–50.

Clarke, J.M., Zaidel, E., 1994. Anatomical–behavioral relationships: corpus callosum morphometry and hemispheric specialization. Behavioral Brain Research 64, 185–202.

Conover W. J.. 1980. Practical Nonparametric statistics, 2nd Edition, John Wiley & Sons, pp. 225-226,

Cootes, T., Edwards, GJ., Taylor, CJ., 2001. Active appearance models. IEEE TPAMI, 23(6),681–685

Cristinacce, D., Cootes, T., Scott, I., 2004. A Multi-Stage Approach to Facial Feature Detection. BMVC04, 231—240

Dager SR., Wang L.. 2007. Shape mapping of the hippocampus in young children with autism spectrum disorder, American Journal of Neuroradiology 28:672-677

Dalal P., Munsell BC., Wang S., Tang J., Oliver K., 2007. A fast 3D correspondence method for statistical shape modeling, IEEE Conference on Computer Vision and Pattern Recognition

Dalal P., Ju L., McLaughlin M., Zhou X., Fujita H., Wang S.. 2009. 3D Open-Surface Shape Correspondence for Statistical Shape Modeling: Identifying Topologically Consistent Landmarks, ICCV

Dalal, AB., Phadk, SR., 2007. Morphometric analysis of face in dysmorphology. Computer Methods and Programs in Biomedicine, 85(2), 165-172

Darken, C.J. & Moody, J. 1990, Fast adaptive kmeans clustering: some empirical results, in 'Proceedings International Conference on Neural Networks, New York: IEEE Computer Science Press', Vol.2, pp. 233-238.

Davies, R.H., Twining, C.J, Cootes, T.F, Waterton, J. C, Taylor, C.J. 2002. A Minimum Description Length Approach to Statistical Shape Modelling, IEEE Trans Med. Imaging, Vol 21 () 525-537.

Dekaban, A. S. 1978. Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. Annals of Neurology, 4(4), 345–356.

Dey T., Giesen J., 2001. Detecting undersampling in surface reconstruction. Proc. 17th ACM Sympos. Comput. Geom., 257-263

Ding, L., Martinez, A., 2008. Precise detailed detection of faces and facial features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.

Dryden, I.L., Mardia, K.V., 1998. Statistical shape analysis. John Wiley and Sons, New York.

Duan Y., Hua J. Qin H.. 2005. Interactive shape modeling using lagrangian surface flow. The Visual Computer, 21(5): 279 – 288

Duda, R., Hart, P., 1972. Use of the hough transform to detect lines and curves in pictures. Communication of the Association of Computer Machinery 15(1), 11–15.

Dyer CA. 2000. Comments on the neuropathology of phenylketonuria. *Eur J Pediatr*;159(suppl 2):S107–S108

Dyer CA, Kendler A, Philibotte T, et al. 1996. Evidence for central nervous system glial cell plasticity in phenylketonuria. J Neuropathol Exp Neurol,55:795–814

Ercan, I., Ocakoglu, G., Guney, I., Yazici, B., 2008. Adaptation of Generalizability Theory for Inter-Rater Reliability for Landmark Localization. International Journal of Tomography & Statistics 2008, Vol. 9, No. S08, 51-58.

Eveno, N., Caplier, A., Coulon, PY., 2001. A New color transformation for lips segmentation. Proceedings of IEEE Fourth Workshop on Multimedia Signal Processing, 3–8.

Eveno, N., Caplier, A., Coulon, PY., 2002. Key Points Based Segmentation of Lips, Proc. ICME'02

Eveno, N.; Caplier, A.; Coulon, PY., 2004. Accurate and quasi-automatic lip tracking. IEEE Transactions on Circuits and Systems for Video Technology,14(5), 706- 715

Farkas L.. 1994. Anthropometry of the Head and Face. Raven Press.

Farkas L., Posnick J.. 1992. Growth and Development of Regional Units in the Head and Face Based on Anthropometric Measurements. The Cleft Palate-Craniofacial Journal, 29(4)

Feris, RS., Gemmell, J., Toyama, K., Krüger, V., 2001. Hierarchical Wavelet Networks for Facial Feature Localization. ICCV'01 Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems

Ferrarinia, L., Palmb, W. M., Olofsena, H., Van Buchemb, M. A., Reibera, J. H.C., Admiraal-Behloula, F.. 2006 . Shape differences of the brain ventricles in Alzheimer's disease, NeuroImage, 32(3), 1060-1069

Fink, W.L., Zelditch, M.L., 1995. Phylogenetic analysis of ontogenetic shape transformations: a reassessment of the piranha genus pygocentrus (teleostei). System Biology, 44(3), 343-360.

Franti, P., Xu, M., 2004. Iterative K-Means Algorithm Based on Fisher Discriminant Fusion: Seventh International Conference on Information Fusion; Stockholm; Sweden

Frazier TW, Hardan AY. 2009. A Meta-Analysis of the Corpus Callosum in Autism. Biol Psychiatry. Sep 11.

Giedd J. N., Castellanos F. X., Rajapakse J. C., Vaituzis A. C., Rapoport J. L., 1997. Sexual dimorphism of the developing human brain, Prog Neuropsychopharmacol Biol Psychiatry, 21(8), 1185-1201

Gibou F., Fedkiw R.. 2005. A fast hybrid k-means level set algorithm for segmentation. In 4th Annual Hawaii International Conference on Statistics and Mathematics, pages 281–291

Gourier, N., Hall, D., Crowley, JL., 2004. Facial features detection robust to pose, illumination and identity, International Conference on Systems Man and Cybernetics, 617–622.

Gourovitch ML, Craft S, Dowton SB, Ambrose P, Sparta S. 1994. Interhemispheric transfer in children with early-treated phenylketonuria. J Clin Exp Neuropsychol, 16(3):393-404.

Gower J. C. 1975. Generalized procrustes analysis. Psychometrika, Volume 40, Number 1: 33-51

Gu X., Wang Y., Chan T., Thompson P.M., Yau S.. 2004. Genus Zero Surface Conformal Mapping and Its Application to Brain Surface Mapping. IEEE Transactions on Medical Imaging, 23(8)

Gu X., Yau S., 2003. Global conformal surface parameterization, in Proc. ACM Symp. Geometry Processing: 127–137

Hahnel S.. 2008. Brain MRI abnormalities in Phenylketonuria, *Clinical Neuroradiology*, 18: pp. 19-24

Hammer, Ø., Harper, D.A.T., Ryan P.D., 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4(1): 9

Hammond P.. 2007. The use of 3D face shape modelling in dysmorphology. Arch. Dis. Child., 92(12): 1120 - 1126.

Hammond P., Hutton TJ., Allanson JE., Buxton B., Campbell LE., Clayton-Smith J., Donnai D., Karmiloff-Smith A., Metcalfe K., Murphy KC., Patton M., Pober B., Prescott K., Scambler P., Shaw A., Smith ACM., Stevens AF., Temple IK., Hennekam R., Tassabehji M., 2005. Discriminating Power of Localized Three-Dimensional Facial Morphology, Am J Hum Genet. 77(6): 999–1010.

Hardan A., Minshew N.J., Keshavan M.S., 2000. Corpus callosum size in autism, Neurology, 55, pp. 1033–1036

Hardan A., Girgis R., Adams J., Gilbert A., Melhem N.M., Keshavan M., Minshew N.. 2007. Brief Report: Abnormal Association Between the Thalamus and Brain Size in Asperger's Disorder. J Autism Dev Disord

Hardan A, Pabalan M, Gupta N, Bansal R, Melhem NM, Fedorov S, Keshavan MS, Minshew NJ. 2009. Corpus callosum volume in children with autism. Psychiatry Res. 174(1):57-61.

Hamarneh G., McIntosh C.. 2005. Physics-based deformable organisms for medical image analysis. SPIE Medical Imaging, 326—335

Hasselbalch S., Knudsen G. M., Toft P. B., et al. 1996. Cerebral glucose metabolism is decreased in white matter changes in patients with phenylketonuria, *Pediatr Res*, 40: pp. 21-24.

Hata Y*., Kobashi S*., Hirano S*., Kitagaki H., Mori E.. 2000. Automated Segmentation of Human Brain MR Images Aided by Fuzzy Information Granulation and Fuzzy Inference, IEEE Transactions on Systems, Man, and Cybernetics—PART C: APPLICATIONS AND REVIEWS, VOL. 30, NO. 3

Heckenberg G., Xi Y., Duan Y., Hua J.. 2006. Brain Structure Segmentation from MRI by Geometric Surface Flow. International Journal of Biomedical Imaging, 6 pages

Hennessy RJ., McLearie S., Kinsella A., Waddington JL., 2005. Facial surface analysis by 3D laser scanning and geometric morphometrics in relation to sexual dimorphism in cerebral–craniofacial morphogenesis and cognitive function, *J. Anat.* 207, pp283–295

Hennessy RJ., McLearie S., Kinsella A., Waddington JL.. 2006. Facial Shape and Asymmetry by Three-Dimensional Laser Surface Scanning Covary With Cognition in a Sexually Dimorphic Manner. J Neuropsychiatry Clin Neurosci 18:73-80

Hochberg, Y., Benjamini, Y., 1995. Controlling false discovery rate: A practical and powerful approach to multiple testing, Journal of the Royal Statistical Society*: Series B, no. 57, pp. 289–300.

Hojjatoleslami S.A., Kittler J., 1995. Region Growing: A New Approach, IEEE Transactions on Image Processing

Huang, F.Z. Su, J., 2002. Face contour detection using geometric active contours, In: Proceedings of the Fourth World Congress on Intelligent Control and Automation, pp. 2090-2093

Jin M., Wang Y., Yau S. T., Gu X., 2004. Optimal Global Conformal Surface Parameterization, *Proc. Visualization,* pp.267-274

Jolliffe, I.T. 1986. Principal Component Analysis. Springer-Verlag.

Jones T., Metaxas D.,1997a. Automated 3D segmentation using deformable models and fuzzy affinity, IPMI 97

Jones T.., Metaxas D..1997b Segmentation Using Deformable Models With Affinity-Based Localization, CVRMed

Jones T.., Metaxas D. 1998. Image Segmentation based on the Integration of Pixel Affinity and Deformable Models. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p.330

Jost J., 2000. Compact Riemann Surfaces, Springer,

Kampmann, M., Zhang, L., 1998. Estimation of eye, eyebrow and nose features in videophone sequences. Int. Workshop on Very Low Bitrate Video Coding, 101–104

Kass. M., Witkin A., Terzopoulos. D. 1988. Snakes: active contour models. International Journal of Computer Vision, 1(4): 321–331,

Kampmann, M., 2004. MAP estimation of chin and cheek contours in video sequences. EURASIP J. Appl. Signal Process., 913-922.

Karsch K., Grinstead B., He Q., Duan Y. 2008. Web Based Brain Volume Calculation for Magnetic Resonance Images, 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

Keary CJ, Minshew NJ, Bansal R, Goradia D, Fedorov S, Keshavan MS, Hardan AY. 2009. Corpus callosum volume and neurocognition in autism. J Autism Dev Disord, 39(6):834-41.

Klauck SM. 2006 . Genetics of autism spectrum disorder. Eur J Hum Genet. 14(6):714-20.

Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, Hamer RM, Lin W, Gerig G, Gilmore JH. 2008. A structural MRI study of human brain development from birth to 2 years. J Neurosci. Nov 19;28(47):12176-82.

Lam KM., Yan H., 1998. An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.20 n.7, p.673-686

Lapeer, R.J.A., Prager, R.W., 2000. 3D shape recovery of a newborn skull using thin-plate splines. Computerized Medical Imaging and Graphics 24, 193-204.

Lee C., Huh S., Ketter TA., Unser M.. 2000. Automated segmentation of the corpus callosum in midsagittal brain magnetic resonance images. Optical Engineering, 39(4): 924-935

Lefohn A. E., Kniss J. M., Hansen C. D., Whitaker R. T.. 2004. A streaming narrow-band algorithm: Interactive computation and visualization of level sets. IEEE Transactions on Visualization and Computer Graphics, 10(4):422–433

Liang J., McInerney T., Terzopoulos D.. 2006. United snakes. Medical Image Analysis, 10: 215–233

Lievin, M., Delmas, P., Coulon, PY., Luthon, F., Fristol, V., 1999. Automatic lip tracking: Bayesian segmentation and active contours in a cooperative scheme. IEEE International Conference on Multimedia Computing and Systems, 1, 691-696

Loizou C. P., Pattichis C. S., Pantziaris M., Tyllis T., Nicolaides A., 2007. Snakes based segmentation of the common carotid artery intima media, Medical and Biological Engineering and Computing, Volume 45, Number 1

Lintas C, Persico AM. 2009. Autistic phenotypes and genetic testing: state-of-the-art for the clinical geneticist. J Med Genet. 46(1):1-8.

Loos, HS., Wieczorek, D., Würtz, RP., Malsburg, C., Horsthemke, B., 2003. Computer-based recognition of dysmorphic faces. Eur J Hum Genet. 11(8), 555-60

Losh M, Sullivan PF, Trembath D, Piven J. 2008. Current developments in the genetics of autism: from phenome to genome. J Neuropathol Exp Neurol. 67(9):829-37.

Lorensen W.E., Cline H.E. 1987. Marching Cubes: A High Resolution 3D Surface Construction Algorithm, Proceedings of SIGGRAPH '87, Vol. 21, No. 4, pp. 163-169

Lucey S., Sridharan S., Chandran V., 2002. Adaptive mouth segmentation using chromatic features, in Pattern Recognition Letters, vol. 23, pp. 1293-1302

Malladi R., Sethian J., Vermuri B., 1993. A topology independent shape modeling scheme. Proc. SPIE Conf. Geometric Methods Comput. Vis. II 2031, 246-258

McInerney T., Akhavan Sharif M.R.. 2006. Sketch initialized snakes for rapid, accurate and repeatable interactive medical image segmentation. 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 398-401

McInerney T., Hamarneh G., Shenton M., Terzopoulos D.. 2002. Deformable organisms for automatic medical image analysis. Medical Image Analysis, 6: 251-266,

McInerney T., Terzopoulos D.. 1996. Deformable models in medical image analysis: a survey. Medical Imaging Analysis, 1(2): 91-108

Miles J.H., Takahashi T.N., Bagby S., Sahota P.K., Vaslow D.F., Wang C.H., Hillman R.E., Farmer J.E., 2005. Essential vs Complex Autism: Definition of Fundamental Prognostic Subtypes, Am. J. Medical Genetics, 135A, pp.171-180

Miroue, M. A., ROSENBERG, L. 1975. The human facial sutures; a morphological and histological study of age changes from 20 to 95 years. M.S.D. thesis, University of Washington.

Moses P., Courchesne E., Stiles J., Trauner D., Egaas B., Edwards E., 2000. Regional Size Reduction in the Human Corpus Callosum Following Pre- and Perinatal Brain Injury. Cerebral Cortex, 10 (12), 1200-1210

Muhle R, Trentacoste SV, Rapin I. 2004. The genetics of autism. Pediatrics. 113(5):e472-86

Mutsvangwa T., Douglas T.. 2007. Morphometric analysis of facial landmark data to characterize the facial phenotype associated with fetal alcohol syndrome. J Anat. February; 210(2): 209–220.

National Institutes of Health Consensus Development Panel. 2000. National Institutes of Health Consensus Development Conference Statement: Phenylketonuria: Screening and management, Pediatrics 2001;108;972-982

Nopoulos, P., Flaum, M. and Andreasen N. C., 1997. Sex Differences in Brain Morphology in Schizophrenia. Am J Psychiatry, 154, 1648–1654

O'Brien L. M., Ziegler D. A., Deutsch C. K., Kennedy D. N., Goldstein J. M., Seidman L. J., Hodge S., Makris N., Caviness V., Frazier J. A., Herbert M. R.. 2006. Adjustment for whole brain and cranial size in volumetric brain studies: a review of common adjustment factors and statistical methods, Harvard Review of Psychiary, 14(3): 141-151,

Oguz I, Cates J, Fletcher T, Whitaker R, Cool D, Aylward S, Styner M. Cortical correspondence using entropy-based particle systems and local features. 2008. Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro: 1637-1640.

Osher S., Sethian J., 1988. Fronts propagating curvature dependent speed: algorithms based on Hamilton-Jacobi formulation. J. Comput. Phy. 79: 12-49,

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Sys., Man., Cyber. **9**, 62–66

Ozdemir ST, Ercan I, Sevinc O, Guney I, Ocakoglu G, Aslan E, Barut C. 2007. Statistical shape analysis of differences in the shape of the corpus callosum between genders, the anatomical record, 29

Pantazis D., Leahy R.M., Nichol T.E., Styner M., 2004. Statistical surface-based morphometry using a non-parametric approach, Int. Symposium on Biomedical Imaging, pp. 1283–1286. 2.1, 2.6

135

Palmer SL., Reddick WE., Glass JO., Gajjar A., Goloubeva O., Mulhern RK.. 2002. Decline in corpus callosum volume among pediatric patients with medulloblastoma: longitudinal MR imaging study. AJNR Am J Neuroradiol, 23: 1088–1094

Pedrycz W. 1993. Fuzzy Control and Fuzzy Systems. London, U.K.: Research Studies Press Ltd.

Pfaendner N. H., Reuner G., Pietz J., Jost G., Rating D., Magnotta V. A., Mohr A., Kress B., Sartor K., Hahnel S.. 2005. MR imaging-based volumetry in patients with early-treated phenylketonuria, *American journal of neuroradiology*, 26(7): pp. 1681-1685

Phillips M. D., McGraw P., Lowe M. J., Mathews V. P., Hainline B. E.. 2001.Diffusion-Weighted Imaging of White Matter Abnormalities in Patients with Phenylketonuria, *Am. J. Neuroradiol.*, 22: pp. 1583-1586

Phillips, P.J., Wechsler, H., Huang, J., Rauss, P., 1998. The FERET database and evaluation procedure for face recognition algorithms. Image and Vision Computing J, 16(5), 295-306

Phillips, P.J., Moon,H., Rizvi, SA., Rauss, PJ., 2000. The FERET Evaluation Methodology for Face Recognition Algorithms, IEEE Trans. Pattern Analysis and Machine Intelligence, 22, 1090-1104

Piven J., Bailey J., Ranson BJ., Arndt S.. 1997. An MRI study of the corpus callosum in autism. Am J Psychiatry 154: 1051-1056

Radeva, P.; Serrat, J.; Marti, E, 1995. A snake for model-based segmentation, Fifth International Conference on Computer Vision

Richtsmeier, J., Paik, C., Elfert, P., Cole, T., Dahlman, H., 1995. Precision, repeatability and validation of the localization of cranial landmarks using computed tomography scans. The Cleft Palate-Craniofacial Journal 32, 217-227.

Ronfard R.. 1994. Region-based strategies for active contour models. Intl. J. of Computer Vision, 13(2):229–251

Ronsen C., Denijver P.A., 1984. Connected components in binary images; the detection problem. Research Studies Press

Rosas, A., Bastir, M., 2002. Thin-plate analysis of allometry and sexual dimorphism in the human craniofacial complex. American Journal of Physical Anthropology 117, 236-245.

Schaefer GB, Mendelsohn NJ; Professional Practice and Guidelines Committee. 2008. Clinical genetics evaluation in identifying the etiology of autism spectrum disorders. Genet Med, 10(4):301-5.

Schoen R., Yau S. T., 1997. Lectures on Harmonic Maps, International Press

Scriver CR, Kaufman S, Eisensmith RC, Woo SLC. 1995. The hyperphenylaninemias. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Metabolic basis of inherited disease. Vol 1. 17th ed. New York, NY: McGraw-Hill;1015–107

Shaffer, J. P., 1995. Multiple Hypothesis Testing. Annual Review of Psychology 46, 561-584.

Sohail, A., Bhattacharya, P., 2007. Automated Lip Contour Detection Using the Level Set Segmentation Method. 14th International Conference on Image Analysis and Processing

Smith S.M., Jenkinson M. , Woolrich M.W., Beckmann C.F., Behrens T.E.J., Johansen-Berg H. , Bannister P.R., De Luca M. , Drobnjak I. , Flitney D.E., Niazy R. , Saunders J. , Vickers J. , Zhang Y. , De Stefano

N. , Brady J.M., Matthews P.M.. 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23(S1):208-219

Stalling D., Hege H.. 1996. Intelligent scissors for medical image segmentation. In B. Arnolds, H. Müller, D. Saupe, T. Tolxdorff (eds). Tagungsband zum 4. Freiburger Workshop "Digitale Bildverarbeitung in der Medizin", Freiburg, März 32-36

Styner M., Oguz L., Xu S., Brechbuehler C., Pantazis D., Levitt J.J., Shenton M.E., Gerig G., 2006. Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM, ISC/NA-MIC Workshop on Open Science at MICCAI

Styner, M., Lieberman, J. A., McClure, R. K., Weinberger, D. R., Jones, D.W. and Gerig, G., 2005. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors, Proc. Natl. Acad. Sci. U. S. A., **102**, 4872–4877

Sun D., Wu L., 2002. Face boundary extraction by statistical constraint active contour model, in: Proceedings of the International Conference on Systems, Man and Cybernetics, vol. 6, Hammamet, Tunisia, pp.14-17.

Tanaka, K., Sano, M., Ohara, S., Okudaira, M., 2000. A parametric template method and its application to robust matching. IEEE Conference on Computer Vision and Pattern Recognition, 1, 620–627.

Tang C.K., Medioni G., Lee M.S., 2000. Tensor Voting, in Perceptual Organization for Artificial Vision Systems, Kim Boyer and Sudeep Sarkar (Eds), Kluwer Academic Publishers, Boston

Theodore, M.C. III, Richtsmeier, J.T., 1998. A simple method for visualization of influential landmarks when using Euclidean distance matrix analysis. American journal of physical anthropology 107, 273–283.

Thompson A. J., Tillotson S., Smith I., et al. 1993. Brain MRI Changes in Phenylketonuria: Associations with Dietary Status, *Brain*, 116: pp. 811-821

Thompson, P. M., Hayashi, K. M., de Zubicaray, G. I., Janke, A. L., Rose, S. E., Semple, J., Hong, M. S., Herman, D. H., Gravano, D., Doddrell, D. M. and Toga, A. W., 2004. Mapping hippocampal and ventricular change in Alzheimer disease, Neuroimage, 22(4), 1754-1766

Tsatsanis KD, Rourke BP, Klin A, Volkmar FR, Cicchetti D, Schultz RT. 2003, Reduced thalamic volume in high-functioning individuals with autism. Biol Psychiatry. 15;53(2):121-9.

Vezhnevets, V., Degtiareva, A., 2003. Robust and Accurate Eye Contour Extraction. Proc. Graphicon-2003, 81-84

Vidal C., Nicolson R., DeVito T., Hayashi K., Geaga J., Drost D., Williamson P., Rajakumar N., Sui Y., Dutton R., Toga A., Thompson P.. 2006. Mapping corpus callosum deficits in autism: an index of aberrant cortical connectivity. Biological Psychiatry, 60(3):218-225

Viola P., Jones M., 2001. Rapid Object Detection using a Boosted Cascade of Simple Features, CVPR, pp. I. 511-518, 2

Viola P., Jones M., 2004. Robust real-time face detection, International Journal of Computer Vision 57 (2):137–154.

Wakasugi, T., Nishiura, M., Fukui, K., 2004. Robust lip contour extraction using separability of multi-dimensional distributions. Proc. 6thIEEE Int. Conf. Automatic Face and Gesture Recognition, 415–420.

Wang, J., Su, G., The research of chin contour in fronto-parallel images. 2003. In: Proceedings of the International Conference on Machine Learning and Cybernetics, pp. 2814-2819.

Wang, SL., Leung, SH., Lau, WH., 2002. Lip segmentation by fuzzy clustering incorporating with shape function. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1, 1077–1080.

Wang, SL., Laua, W.H., Leung, S.H., 2004. Automatic lip contour extraction from color images, Pattern Recognition 37, 2375 – 2387

Wang L., Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, Csernansky JG. 2003. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. Neuroimage, 20(2):667-82.

Wang S., Chung MK. 2005. Parametrization and classification of closed anatomical curves. Technical Report, no.1113,

Westin CF., Lorigo LM., Faugeras O., Grimson WEL., Dawson S., Norbash A., Kikinis R.. 2000. Segmentation by Adaptive Geodesic Active Contours. Third International Conference on Medical Image Computing and Computer-Assisted Intervention, Pages 266-275

Wolf H., Kruggel F., Hensel A., Wahlund L., Arendt T., Gertz H.. 2003. The relationship between head size and intracranial volume in elderly subjects, *Brain Research*, 972(1): pp. 74-80

Wood ZJ., Schröder P., Breen D., Desbrun M., 2000. Semi-regular mesh extraction from volumes, Proceedings of the conference on Visualization '00, p.275-282,

Wo¨rz, S., Rohr, K., 2006. Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. Medical Image Analysis, 10, 41–58

Xu M., 2004. Delta-MSE Dissimilarity in GLA-based Vector Quantization, IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, (ICASSP'04), Montreal, Canada

Yokogawa Y., Funabiki, N., Higashino, T., Oda, M., Mori, Y., 2007. A Proposal of Improved Lip Contour Extraction Method Using Deformable Template Matching and Its Application to Dental Treatment. Systems and Computers in Japan, 38(5)

Yotsutsuji, T., Saitoh, O., Suzuki, M.,, Hagino, H., Mori, K., Takahashi, T., Kurokawa, K., Matsui, M., Seto, H. and Kurachi, M., 2003. Quantification of lateral ventricular subdivisions in schizophrenia by high-resolution three-dimensional magnetic resonance imaging, Psychiatry Research: Neuroimaging, 122, 1-12

Yuille, A., Hallinan, P., Cohen, D., 1992. Feature extraction from faces using deformable templates. International Journal of ComputerVision 8, 2, 99–111.

Yushkevich PA., Piven J., Hazlett HC., Smith RG., Ho S.. 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage 31(3): 1116-1128,

Zadeh L.A.. 1997. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets Syst., vol. 90, pp. 111–127

Zadeh L. A. 1965. Fuzzy sets. Information and Control, 8 (3): 338–353.

Zhao H.. 2005. A fast sweeping method for eikonal equations. Mathematics of Computation, 74:603–627, 4

Zheng, Z., Yang, J., Yang, L., 2005. A robust method for eye features extraction on color image, Pattern Recognition Letters, 26, 2252–2261

VITA

Qing He was born on January 18, 1983 in Nanjing, China. She received her Bachelor of Engineering degree in Control Theory from Nanjing University of Science and Technology in 2004, and received her Master of Engineering in Pattern Recognition and Intelligence System from Shanghai Jiaotong University in 2006. She was admitted by the department of Computer Science, University of Missouri in the fall of 2006 and began to pursue her PhD degree. She has worked for Dr Ye Duan as a research assistant since then. Her research interests are medical image segmentation, statistical shape analysis of brain structures, and facial feature localization.