MEDICAL IMAGE ANALYSIS USING STATISTICAL SHAPE MODEL BASED ON SUBDIVISION SURFACE WAVELET

LI YANG

B. Eng, Xi'an Jiaotong University, P. R. China
M. Eng, Xi'an Jiaotong University, P. R. China

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

NATIONAL UNIVERSITY OF SINGAPORE

2007

Acknowledgements

I would like to express my deepest appreciation to my supervisors, Assoc. Prof. Tan Tiow-Seng, Prof. Wieslaw L. Nowinski and Dr. Ihar Volkau for their expert and enlightening guidance in the achievement of this work. They gave me lots of encouragement and constant support throughout my Ph.D studies, and inspired me to learn more about medical image analysis and other research areas.

I would also like to thank my colleagues and friends in the Biomedical Imaging Lab and the Computer Graphics Research Lab for their generous help and warm friendship during these years.

Finally, I would like to extend my sincere thanks to my family. They have been a constant source of love and support for me all these years.

Contents

Abstract

Statistical shape models which represent the shape variations within a population are used in a variety of applications of medical image analysis, such as model-guided segmentation, statistical shape analysis and probabilistic atlasing. In this thesis, we propose a novel statistical shape model based on the shape representation using subdivision surface wavelets. It has three highly desirable properties of a statistical shape model: compact shape representation, multi-scale shape description and spatial-localization of the shape variation.

We also develop a new model-guided segmentation framework utilizing this Statistical Surface Wavelet Model (SSWM) as a shape prior. In the model building process, a set of training shapes are decomposed through the subdivision surface wavelet scheme. By interpreting the resultant wavelet coefficients as random variables, we compute prior probability distributions of the wavelet coefficients to model the shape variations of the training set at different scales and spatial locations. With this statistical shape model, the segmentation task is formulated as an optimization problem to best fit the statistical shape model with an input image. Due to the localization property of the wavelet shape representation both in scale and space, this multi-dimensional optimization problem can be efficiently solved in a multiscale and spatially localized manner. We have applied our method to segment cerebral caudate nucleus and putamen from MR (Magnetic Resonance) scans of both healthy controls (27 cases) and patients with schizophrenia (38 cases). The experiment results have been validated with manual segmentations. The results show that our segmentation method is robust, computationally efficient and achieves a high degree of segmentation accuracy. After that, a comparative statistical shape analysis of the caudate nucleus between schizophrenia patients and normal controls is performed as well. In the statistical group mean difference hypothesis testing between schizophrenia patients and healthy controls regardless of gender, race and handedness, significant shape difference between the two groups is suggested. In order to exclude the unknown affects of gender, race and handedness to the shape analysis, the same hypothesis testing is also conducted on two sub-groups which only consists of right-handed Chinese male. However, in this test, no significant shape difference between the two groups is clearly suggested. Considering the relative insufficient subjects in this analysis (only 17 schizophrenia patients and 8 healthy controls), a further study based on more datasets is necessary.

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

FLD	Fisher's Linear Discriminant
MAP	maximum a posteriori probability
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSE	mean-square error
PCA	Principal Component Analysis
PDF	probability density function
SNR	signal-to-noise ratio
SPHARM	Spherical Harmonics
SSA	Statistical Shape Analysis
SSM	Statistical Shape Model
SSWM	Statistical Surface Wavelet Model
SVM	Support Vector Machines

Symbols

$(\cdot)^\top$	the transpose operation
$\det(\mathbf{A})$	the determinant of matrix $\mathbf{A}$

$\mathbb{N}$	the natural number field
$\mathcal{N}(u, \sigma^2)$	the Gaussian distribution with mean u and variance σ^2
$\exp(\cdot)$	the exponential function
$\Pr(\cdot)$	the probability of the event in the brackets
$\mathrm{Re}\{\cdot\}$	the real part of the quantity in the brackets

Chapter 1

Introduction

There exists of a large number of objects with different shapes — from the non-life-form: planets, molecular and atom to the life-form: anatomical structures, tissues and cells. The shape of an object lies at the interface between vision and cognition [1]. Therefore, the analysis of the shape is usually the first step we take to get a profound understanding of the objects we are investigating. This is especially true in biology and medical researches, because the shape or shape variations of anatomical structures is closely related to their physiological functions. The branch that deals with the shape and structure of organisms has become an important sub-domain, *Morphology*. The morphology study in medicine is usually based on the biomedical images generated from CT, MR scan, X-ray, PET, etc. Originally, only simple measurement of size, area, volume, orientation and symmetry of the individual anatomical structures are used. However, the changes in these metrics are only general features, because although they might explain the atrophy or dilation caused by illness, the morphological changes at specific locations are not sufficiently reflected in these global metrics. Therefore, full geometrical information, especially the local shape information should be taken into account. At the same time, the analysis based on a single object can't answer some of the key questions in medical morphology study. For example, what is the shape of the human brain surface? It is quite difficult to answer this question, because the shape differs from person to person. Instead of giving a single and fixed brain surface atlas, it is much more appropriate to give a probabilistic brain surface atlas which indicates the different variation modes at different locations among different groups of peoples. Another example is how to discriminate between the normal morphological variations of brain structures and the pathological variations caused by neurological diseases, for instance, schizophrenia? The answers can only come from the statistical and quantitative comparison and analysis between healthy and diseased subjects. Thus, *Statistical Shape Analysis* (SSA), which no longer analyzes only single or several subjects but a quite large population, has become of increasing interest to the medical imaging community.

1.1 Statistical Shape Analysis (SSA) and Statistical Shape Model (SSM)

Given a population, there are generally pronounced anatomical variations among the subjects. Statistical shape analysis of medical images aims to study the various statistical quantities of these variations. The main objective of statistical shape analysis is to provide a probabilistic description of shapes, a quantitative measurement of shape variation and a classification of shapes according to the variation mode. Statistical shape analysis is, therefore, potentially capable of precisely locating the pathological variations or understanding and quantifying how the factors, such as diseases, aging, gender and races *et al.*, affect these morphological changes. However, before statistical shape analysis can be conducted, it is necessary to have a standard shape description in which shapes from different subjects are comparable and a framework to perform the comparison and statistical analysis of the shape variance. Thus, such a mathematical framework, the so-called *Statistical Shape Model* (SSM), which provides these essentials, is the pivotal problem in statistical shape analysis. Statistical shape analysis, in fact, can be regarded as a process of building the statistical shape model.

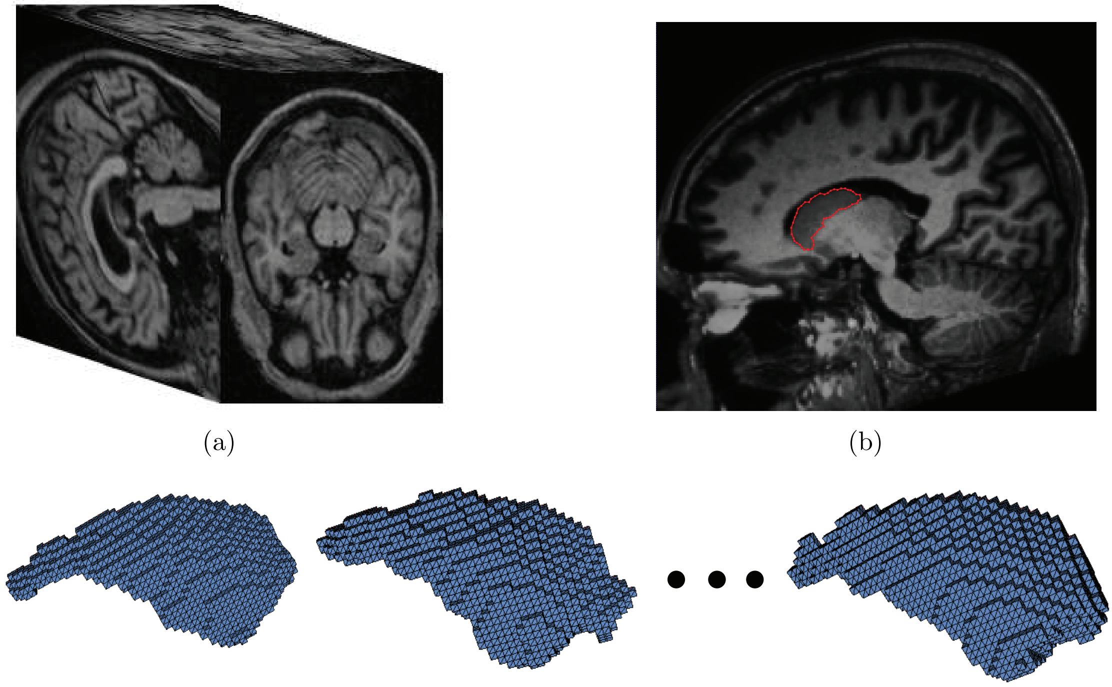
Usually, there are 3 major steps in the construction of a statistical shape model (or performing statistical shape analysis): (1) image data preparation; (2) shape representation; (3) statistical analysis. In the remaining part of this section, we will give a brief introduction to these main steps.

1.1.1 Image Data Preparation

The construction of a statistical shape model starts from the data collection and preparation. Firstly, a considerable number of planar or volumetric scans for the subject in the study are acquired. Then, the anatomical structures of interest are segmented, either manually or using automatic algorithms designed for this task [2–7]. As an example, Fig. 1.1(a) shows one 3D MR scan of a human head. Fig. 1.1(b) shows the manual segmentation of the caudate nucleus in a sagittal slice. Fig. 1.1(c) shows several finally segmented caudate nucleus from scans of different subjects. Although the segmented examples in Fig. 1.1(c) look very similar, their shape and volume difference can be detected even through the visual inspection. However, far beyond the detection of the differences, the goal in the morphology and pathology studies is to not only localize morphological differences using shape information but also to quantify them for assessing the severity of the disorder, effectiveness of the treatment or correlating them with symptoms. To achieve the quantitative analysis of morphological variations, shapes of different subject must be compared between each other. Therefore, instead of representing the shape in voxels, an uniform shape representation is needed, in which different shapes are comparable.

1.1.2 Shape Representation

Usually, tens, hundreds, or even thousands of examples are analyzed in the statistical shape analysis or used to build a statistical shape model. Therefore, in order to make the shape of different objects comparable, the segmented objects are transformed into other shape representations (usually in vector form).



(a)

(b)

(c)

Fig. 1.1: Data preparation. (a) The volumetric MR image. (b) The manual segmentation of caudate nucleus in one sagittal slice. (c) Three segmented caudate nucleus in volumetric binary image form (from IBSR [8]).

A great number of shape descriptors have been proposed over the years for this purpose. For example, the simplest and straightforward method is to represent the shape by the same number of sample points on the object boundary [2, 9]. Another approach is to describe the object boundary through modal decomposition [3, 5, 10, 11]. Different from the shape representation methods by direct outlining the object boundary, deformation fields [6, 12, 13] were also used as shape representations in building statistical shape models. The shape representation is the crucial problem in statistical shape model building, because the property of the shape representation directly determines the capability and properties of the resulting statistical model and statistical shape analysis. A detailed comparison of the existing shape representation methods used in statistical shape models is given in Chapter 2.

1.1.3 Statistical Analysis

Once the shape of segmented objects have been transformed into the vector-form shape representation, they are now comparable and statistical analysis can be performed on these shape vectors to find the statistical features of a dataset. Generally, this is typically done by applying Principal Component Analysis (PCA) to the dataset. The mean shape vector is then considered a “typical” shape, and the principal components are computed to capture the variation within the set and used to represent new input shape. If there are two or more comparative populations, comparative or discriminative analysis will be conducted to find the shape differences between these populations. Finally, if significant shape difference exist between populations, a shape classifier (FLD [14] or SVM [15, 16]) in the shape vector space can be trained to separate the different populations.

1.2 Statistical Shape Model and Model-Guided Segmentation

In last section, we have explored the 3 main steps of building a statistical shape model or performing statistical shape analysis. It is obvious that the precise segmentation of the object in data preparation is a prerequisite step. The accuracy of the segmentation determines the quality of the subsequent statistical shape model or statistical shape analysis. In fact, segmentation is absolute prerequisite and necessary for a variety of applications: i.e. pre-operative evaluation and surgery planning [17], radiotherapy treatment planning [18] and monitoring of disease progression or remission [19]. While this task has traditionally been tackled by human experts, the drawbacks of manual segmentations, such as time-consuming, lack of reproducibility and subjective biases, make an automatic or semi-automatic method highly desirable. However, because of the highly variable nature of the shapes of anatomical structures, an accurate automated segmentation method is a true challenge. Low level segmentation algorithms (region growing [20], edge detection [21], snake [22]) may be used to assist the human operator, but reliable results could hardly be expected without human intervention because of the many difficulties [3, 23]: input images are noisy (very low SNR (signal-to-noise ratio)), not very well contrasted, surrounding structures with similar shape or intensity, the target structure are fairly variable in shape and intensity, etc.. Therefore, to overcome these difficulties, high level model-guided methods have been proposed [2–4, 10]. In these methods, the statistical shape model was used as probabilistic template to introduce the prior knowledge into the automatic segmentation process. Compared to other fixed shape template/model, the statistical shape model is much more suitable for this purpose, because it contains all the known variations of the structures.

1.3 Thesis Contributions

Although 2D statistical shape model based on the first generation wavelet [11, 24] has been proposed and shown to be a better choice for statistical shape analysis especially for spatial localized shape variations, the rigorous requirements in the explicit surface parameterization required by the first generation wavelets scheme are the main obstacles of the extension of this method to 3D surface.

In order to address this problem, the main purpose of this thesis is to develop a novel statistical shape model for the genus-zero object (the most common topology of biological objects) based on the subdivision surface wavelet transform, termed Statistical Surface Wavelet Model (SSWM). And besides, a framework of using SSWM for model-guided segmentation and comparative shape analysis will be proposed. Our new model adopts a newly developed surface wavelet scheme based on the lifting scheme [25]. This scheme can perform wavelet transforms on irregular grids. Thus, as a result, the SSWM doesn't need the surface to be explicitly parameterized, so that it can perform the shape analysis directly on the surface mesh with certain subdivision connectivity. Therefore, a method to prepare the surface mesh with correspondence in certain mesh-connectivity which is required by the wavelet scheme will also be presented in the thesis.

Because of the adoption of wavelet basis, the SSWM is expected to possess all the following three highly desirable properties simultaneously: compact shape representation, multi-scale shape description, and spatial-localization of the shape variation. These good properties will be advantageous in applications using statistical shape models, such as statistical shape analysis and model-guided segmentation. Firstly, we will use this model to investigate a shape population consisting of 18 caudate nuclei. The model is designed such that shape analysis can be focused on scale and spatial location on the surface. Such a multiscale and spatially localized shape analysis, which is not possible in previous models, can be very useful as diseases, such as cancer, may only affect a small portion of an organ. Furthermore, the resultant multiscale and spatially localized statistical shape model can be used in model-guided segmentation. In the segmentation process, fitting the model to the image is, in general, an optimization problem. However, too many input parameters to an optimizer at a time will lead to extremely high computational cost. In the previous models, because of the lack of spatial localization in shape space, all the model parameters in one scale level have to be inputted together for optimization. In some cases, this even causes the optimization computationally impracticable. In contrast, the SSWM can be fitted with the image in a divide-and-conquer manner. The whole model fitting problem is solved by optimizing the model parameter one by one, since each of them only defines the shape at certain scale and spatial location. This is expected to result in a much more efficient and robust model-guided segmentation method.

1.4 Outline of the Thesis

The remaining part of the thesis is arranged as follows. Firstly, in Chapter 2, a review of related work is presented. The focus is on the existing shape representations and their properties relevant to the problem of statistical shape model and statistical shape analysis. The purpose of this chapter is to provide a general overview of commonly used shape representations, as well as guidelines for choosing a shape representation for the statistical shape model building and model-guided segmentation purpose.

In the following chapters, a new statistical shape model based on the subdivision surface wavelet shape representation will be proposed and applied in model-guided segmentation of the caudate nucleus. Since the whole process is quite complex, to help to put things together, an outline of the thesis is given in Fig. 1.2. The left column indicates the steps in the process. The representative results in the steps are shown in the middle column. The right column indicates the Sections where details of the steps will be addressed. The dotted line indicates the partition between model training and model application.

Chapter 3 explains our choice of the subdivision surface wavelet for shape representations and presents the scale and space localization properties of the resultant Statistical Surface Wavelet Model (SSWM). In this chapter, the other two key problems in statistical shape model building, i.e. establishing correspondence between surfaces of different subjects and surface re-meshing, are covered in Section 3.2. After that, in Section 3.3, as an example, a SSWM depicting the shape variations of the caudate nucleus will be constructed based on 18 MR scans from The Internet Brain Segmentation Repository (IBSR) [8]. Next, in Chapter 4, the acquired SSWM is used in model-guided segmentation as a shape prior. By utilizing the "double localization" property of wavelet basis, a multiscale and spatial localized algorithm is proposed to optimize the model fitting objective function. The segmentation experiments of caudate nucleus and putamen were conducted on the MR images from both the schizophrenia and healthy controls. The results

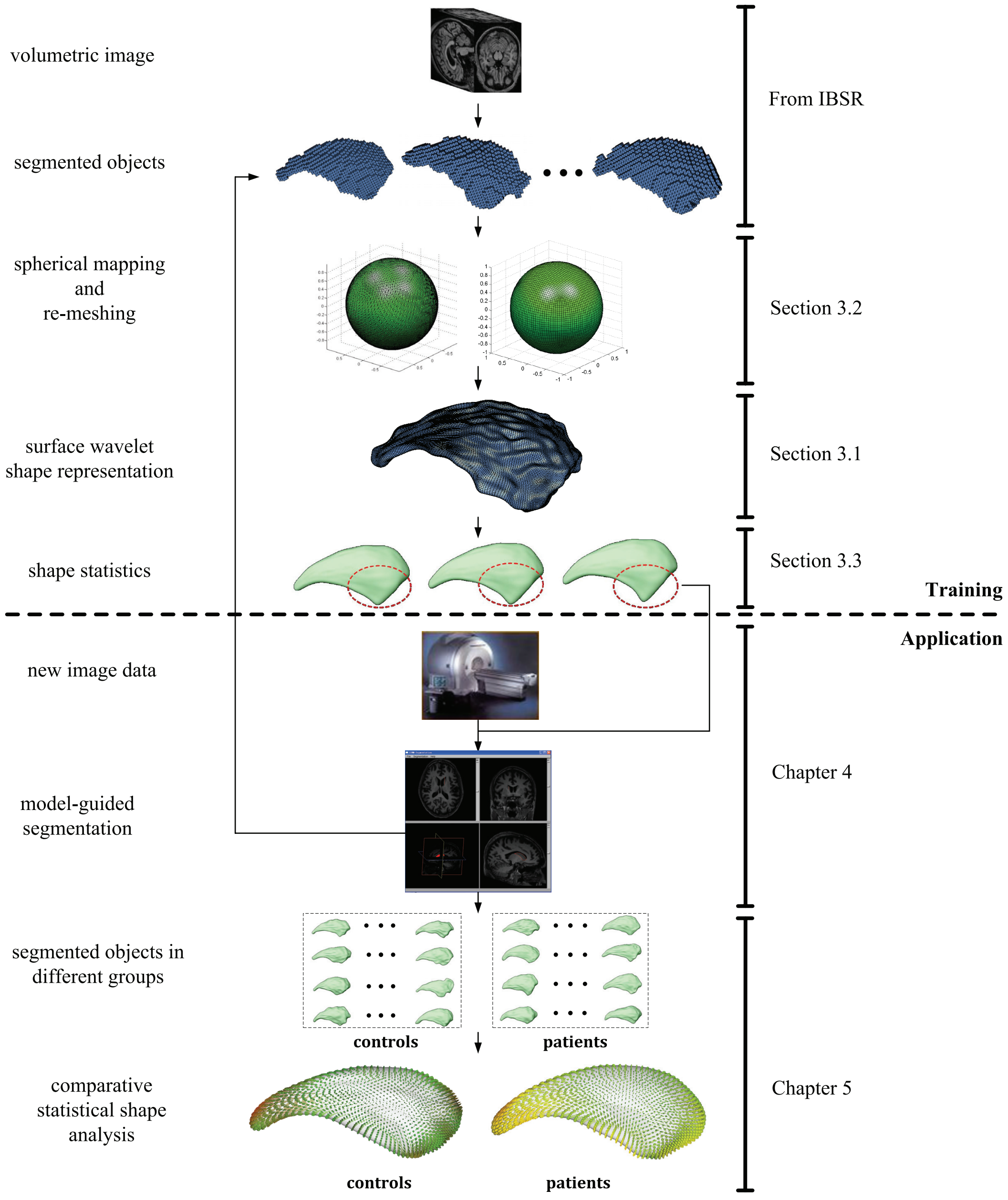


Fig. 1.2: Outline of the thesis

were validated by comparing with the manual segmentations. In Chapter 5, we give the results of comparative shape analysis of caudate nucleus between two groups, schizophrenia patients and healthy controls. In Chapter 6, the thesis concludes with a discussion of the lessons learned from the presented experiments and future research directions enabled by the results of this work.

Chapter 2

Related Work

As mentioned in Section 1.1.2, shape representation (or shape descriptor) is the pivotal problem in building a statistical shape model, performing statistical shape analysis or model-guided segmentation. This chapter reviews the existing shape representations and their relevant properties. We limit our review to include only shape representations that have been used in medical image analysis and model guided-segmentation, while leaving out some shape representations used in computer vision or other applications. The purpose of this chapter is to provide a brief overview of existing shape representations and a necessary background for the discussion on our novel statistical shape model based on the surface wavelet shape representation in next chapter. In order to derive properties that are needed for a shape description suited for shape analysis or building a shape prior for automatic segmentation, selected properties of shape descriptions are investigated as well. This investigation leads to a list of properties that outlines the requirements for an ideal shape description scheme for statistical shape model.

2.1 The Classification of Shape Descriptions

There are a large variety of existing geometric representations of shape as illustrated in Fig. 2.1. Depending on their underlying structure, they can be

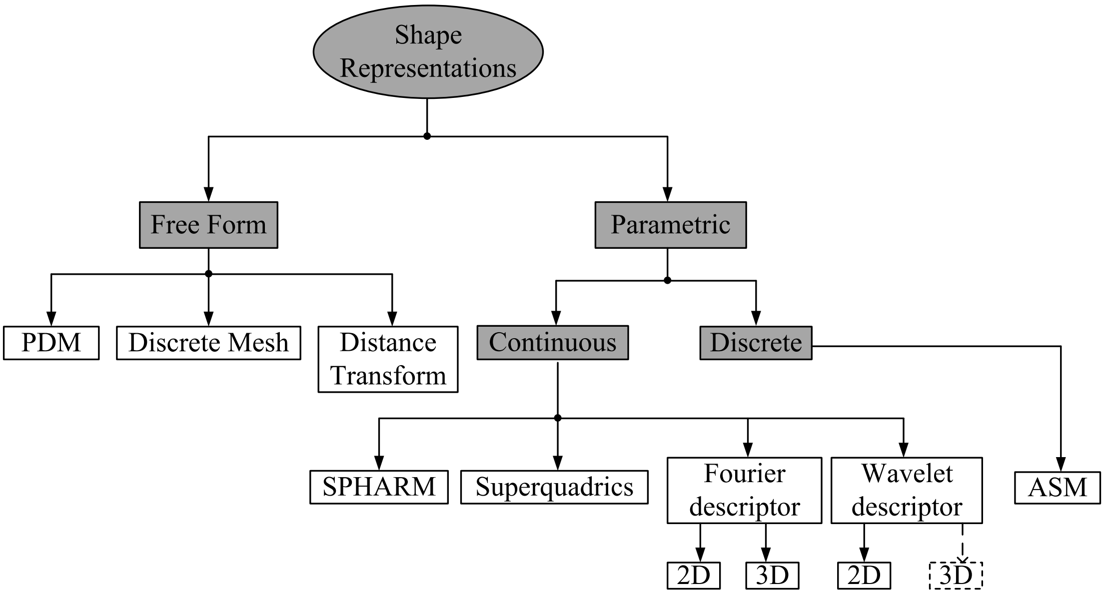


Fig. 2.1: Different geometric representation of shape models

partitioned into 2 classes: free-form and parametric. Both can be used in the construction of statistical shape model and model-guided segmentation, but with different pros and cons. In the next few sections, these different shape representations will be explored by explaining the main idea of the methods.

2.2 Free-Form Shape Descriptions

Free-form shape descriptions are based on explicit or implicit listing of points or patches on the object boundary, which do not assume any specific global structure. The only constraints are local continuity and smoothness. Therefore, they provide considerable flexibility to represent arbitrarily complex shapes. The main drawback of this kind of models is that they are not very concise and lack of overall shape information, because of the use of local primitives (points, facets) on shape boundary.

2.2.1 Point Distribution Model (PDM)

The most representative free-form shape representation method is the Points Distribution Model (PDM), in which the shape is represented by an explicit list of sample points on shape boundary. There are several application of this shape representation in statistical shape analysis, for example, Bookstein in 2D [26–28], Cootes [2, 9, 29] and Rangarajan [30] in 3D. Since only a number of points are selected to represent the shape, shape information between these points is unknown.

Based on this shape representation, an elastic deformable model was introduced by Kass *et al.* [22] to find the boundary of object. In this so-called “snake” method, the shape model deforms from an initial position to fit the edge features in an image. The points on the boundary are represented parametrically as:

$$\mathbf{v}(t) = (x(t), y(t)) \tag{2.1}$$

where the parameter $t \in [0, 1]$ is proportional to the arc-length.

The behavior of the snake is driven by minimization of a cost function that combines image, internal and constraint energies:

$$E = \alpha E_{image} + \beta E_{int} + \gamma E_{con} \tag{2.2}$$

The image energy guides the model to match the edge feature and is derived by integrating over the boundary with an image edge map [31]. The internal energy constrains the model shape to be smooth and is defined as the integral of the first and second order derivative of the boundary, which control the tension and rigidity of the boundary respectively. The constraint energy is introduced to allow user interaction. Later, many modifications of the original snake algorithm were proposed, such as [23, 32]. However, the main drawback of this method is that it is very sensitive to the model initialization.

2.2.2 Discrete Mesh

This method represents the shape as a set of discrete geometric entities, usually triangulation or quadrilateral facets. They are widely used techniques in computer graphics, but are rarely used for shape analysis. The first application of this method in building deformable shape model was done by Delingette who introduced the simplex mesh model [33]. A simplex mesh is a discrete shape representation with a constant vertex connectivity. Each vertex of a 2-simplex mesh is connected to three neighbors. Therefore, 2-simplex meshes are used to represent surfaces. The mesh is adaptive in its density and topology. It has been applied in segmentation by deforming the mesh according to the potential field defined by the object boundary. Approaches for volumetric and shape measurement of simplex mesh have been developed as well. However, since the problem of finding corresponding points on different surfaces remains unsolved for this kind of mesh, no statistical shape analysis work using this model has been done so far.

2.2.3 Distance Transform/Level Set

This method was developed by Osher and Sethian firstly in 2D [34] and has been extended to surfaces [35]. The main idea of level sets is to embed the deformable shape representation in higher dimension space. For example, in 3D, a surface is represented as the zero level set of a function $\Psi : \mathbb{R}^3 \rightarrow \mathbb{R}$:

$$\mathcal{S} = \{\boldsymbol{p} \in \mathbb{R}^3 | \Psi(\boldsymbol{p}) = 0\}$$

Usually, the function Ψ is chosen as distance transform [36], which is a function that for each point in the image is equal to the distance from that point to the boundary of the object. In the signed variant of the original distance transform, the values of the distance transform negates outside the object, in order to eliminate the singularity at the object outline and make the value change linearly when crossing the boundary. In this way, the boundary is modeled implicitly as a zero level-set of the distance transform. The distance transform can be computed from a binary segmentation of the object. The main advantage of level set method is to naturally change the topology of surface (or contour) in the deformation process. The model may split into several components or merge from several components while still remains only one function. Its drawback is the computational cost, since a higher dimension space is used for representing the surface. Moreover, because of its implicit representation of the shape, it is hard to define the correspondence required in building the statistical shape model. Therefore, when it is used for segmentation, the incorporation of prior knowledge is usually through indirect methods, such as intensity statistics [6], inter-objects constraints [37].

2.3 Parametric models

Different from the free-form models using representation in image space, parametric models capture the overall shape in a small number of parameters acquired by decomposing the shape by a set of basis functions or a serial of harmonics with different scales. The shape representation becomes more accurate as higher scale bases are added to the representation. In practice, it is desired to reduce as much as possible the number of bases used in order to obtain a compact representation. However, it is known that a few bases usually allow the representation of rather complex shapes. Therefore, parametric model provides a concise representation of shape. Moreover, the parametric models offer a straightforward way for the inclusion of prior shape knowledge, because probability distributions of the parameters can be easily incorporated to bias the model to a particular overall shape while allowing for deformations in certain degree. More importantly, since they are more concise representations of the shapes, the optimization problem of matching a model to an image data can be solved in a lower dimensional space. Next, several existing parametric shape models will be reviewed.

2.3.1 ASM (Active Shape Model)

This method was developed by Cootes and Taylor [2]. In their work, a shape is still represented by using the point distribution model (PDM). However, beyond that, from a training set, a mean shape is formed by averaging the corresponding points' coordinates (after normalization to exclude size, orientation and position) over all members in the training set. Then, through the principal component analysis on the training set, eigenvectors which represent the eigen shape variation modes are computed and new shapes are modeled by the mean shape plus a linear combination of these eigenvectors. These weights of the linear combination are, in fact, the parameters for shape representation. However, as the training set is usually small in size, relative to the dimensionality of the shape space, the possible ways to deform such a shape model are limited to a linear subspace of the complete shape space. Moreover, the representation of shape is still composed of a set of discrete points. Thus, we know the geometry of a shape only at a finite set of points. When a high degree of precision is required on the shape, a corresponding dense sampling is needed. Thus, this shape model can be verbose and thus not efficient for computation.

2.3.2 Superquadrics

Another approach to parametric surface description is superquadrics, introduced in [38]. The most widely used superquadrics are ellipsoids:

$$\mathbf{v}(a_1, a_2, a_3, a_4, \varepsilon_1, \varepsilon_2, u, v) = a \begin{pmatrix} a_1 C_u^{\varepsilon_1} C_v^{\varepsilon_2} \\ a_2 C_u^{\varepsilon_1} S_v^{\varepsilon_2} \\ a_3 S_u^{\varepsilon_1} \end{pmatrix} \quad \begin{array}{l} -\pi/2 \leq u \leq \pi/2 \\ -\pi \leq v \leq \pi \end{array} \tag{2.3}$$

where $S_\omega^\varepsilon = sgn(\sin\omega)|\sin\omega|^\varepsilon$ and $C_\omega^\varepsilon = sgn(\cos\omega)|\cos\omega|^\varepsilon$ and $a_1, a_2, a_3 \geq 0$ are scale parameters defining the aspect ratios and $\varepsilon_1, \varepsilon_2 \geq 0$ are "squareness" parameters. Fig. 2.2 shows some superquadric ellipsoids for varying squareness

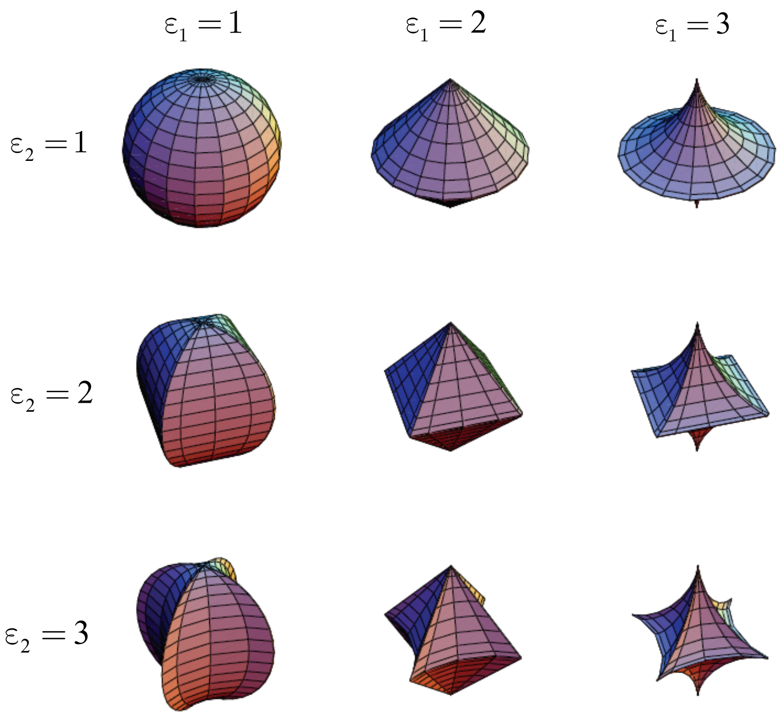


Fig. 2.2: Shapes of superquadric ellipsoids with $a_1, a_2, a_3 = 1$ for different squareness parameters ε_1 and ε_2

parameters.

As we can see in Fig. 2.2, superquadrics alone are not able to represent complex shapes. Therefore, many extensions were proposed. For example, Bardinet *et al.* and Vemuri [39, 40] deform the superquadrics by applying free-from deformations, which bend the original superquadrics to represent more complex shapes. However, because in these extended approaches, the parametric net on the superquadric surfaces is distorted due to the deformation, and therefore, not suited for the approximation of corresponding points on different surfaces. For this reason, superquadrics models are rarely used for building statistical shape model.

2.3.3 Fourier Models

The Fourier model was proposed by Staib and Duncan in 2D [10]. They proposed using a Fourier representation for parameterized deformable contours.

A Fourier representation for a closed contour is expressed as:

$$S(t)=\begin{bmatrix} x(t)\\ y(t)\end{bmatrix}=\begin{bmatrix} a_0\\ c_0\end{bmatrix}+\sum_{k=1}^{\infty}\begin{bmatrix} a_k & b_k\\ c_k & d_k\end{bmatrix}\begin{bmatrix}\cos 2\pi kt\\ \sin 2\pi kt\end{bmatrix} \tag{2.4}$$

where $S(t)$ ($t\in[0,1]$) is a parameterized closed curve in 2D. $x(t)$ and $y(t)$ are the parametric function for x and y coordinates respectively. $a_k, b_k, c_k, d_k, (k=1,2,3,\cdots)$ are Fourier coefficients computed by:

$$\begin{array}{ll} a_0=\frac{1}{2\pi}\int_0^{2\pi}x(t)dt & c_0=\frac{1}{2\pi}\int_0^{2\pi}y(t)dt\\ a_k=\frac{1}{\pi}\int_0^{2\pi}x(t)\cos ktdt & b_k=\frac{1}{\pi}\int_0^{2\pi}x(t)\sin ktdt\\ c_k=\frac{1}{\pi}\int_0^{2\pi}y(t)\cos ktdt & d_k=\frac{1}{\pi}\int_0^{2\pi}y(t)\sin ktdt \end{array} \tag{2.5}$$

These coefficients follow a scale ordering, where low index coefficients describe large scale properties and higher indexed coefficients describe more detailed shape information. After truncating the coefficients series according to the level of detail requirement of specific application, the coefficients can be mapped to form a coefficient vector which can serve as a shape descriptor.

In order to model a group of shapes that have the same topological structures but may differ slightly due to deformation, they interpret the Fourier coefficients as random variables (Gaussian). Therefore, the parameter vector is now a deformable stochastic Fourier shape descriptor. The prior knowledge of the shape is stored in the probability distributions of these parameters in the Fourier descriptor:

$$\Pr(\boldsymbol{p})=\prod_{i=1}^{N}\Pr(p_i)=\prod_{i=1}^{N}\frac{1}{\sigma_{i\sqrt{2\pi}}}e^{-\frac{(p_i-m_i)^2}{2\sigma_i^2}} \tag{2.6}$$

where, $\boldsymbol{p}=[p_1p_2\cdots p_N]$ is the parameter vector of the Fourier shape descriptor. m_i and σ_i are the mean and variance of random variable respectively.

Next, in order to apply the prior knowledge of shape in the process of boundary finding, they formulate the problem using a maximum *posterior* criterion based on Bayes rule. In the boundary finding or segmentation problem, the input is an image $I(x, y)$. t_p is an image template corresponding to a particular value of parameter vector $\boldsymbol{p}$. Therefore, the goal is to locate the object which is depicted by t_p in the image. This decision should be made both on the prior shape information of the object we are looking for and the image information. In terms of probabilities, if we want to decide to which template t_p and image I correspond, we have to evaluate the probability of the template, given the image $\Pr(t_p|I)$, and find the maximum over $\boldsymbol{p}$. This can be expressed using Bayes rule:

$$\Pr(t_{map}|I) = \max_p \Pr(t_p|I) = \max_p \frac{\Pr(I|t_p)\Pr(t_p)}{\Pr(I)} \tag{2.7}$$

where, t_{map} is the maximum of a *posterior* solution, $\Pr(t_p)$ is the prior probability of the template t_p, and $\Pr(I|t_p)$ is the conditional probability of the image given the template. By taking logarithm and eliminating $\Pr(I)$, which is the probability of the image data that will be equal for all $\boldsymbol{p}$. Thus, it suffices to maximize

$$M(I, t_{map}) = \max_p M(I, t_p) = \max_p [\ln \Pr(t_p) + \ln \Pr(I|t_p)] \tag{2.8}$$

This form of objective function shows the trade-off that will be made between prior information $\Pr(t_p)$ and image-derived information $\Pr(I|t_p)$. Till now, we can see that the problem of finding the boundary or segmentation has transformed to an optimization problem of searching in the vector space to find an appropriate $\boldsymbol{p}$ which can maximize Eq. (2.8). Upon the success of using deformable Fourier models in 2D, Staib and Duncan [3] extended this method into 3D. In 3D, a surface can be represented explicitly by 3 functions of 2 index parameters:

$$\mathcal{S}(x, y, z) = (x(u, v), y(u, v), z(u, v)) \tag{2.9}$$

In order to represent surfaces, a basis for functions of 2 variables is used:

$$\begin{aligned} \phi = \{&1, \cos mu, \sin mu, \cos lv, \sin lv, \cos mu \cos lv, \sin mu \cos lv, \\ &\cos mu \sin lv, \sin mu \sin lv, \cdots\} \end{aligned} \tag{2.10}$$

where, $m, l = 1, 2, \cdots$. Each function is then represented by:

$$f(u,v) = \sum_{m=0}^{K_1-1} \sum_{l=0}^{K_2-1} \lambda_{m,l} \left[a_{m,l} \cos mu \cos lv + b_{m,l} \sin mu \cos lv + \right.$$
$$\left. c_{m,l} \cos mu \sin lv + d_{m,l} \sin mu \sin lv + \cdots \right] \tag{2.11}$$

where $\lambda_{m,l} = \begin{cases} 1 \text{ for } m = 0, l = 0 \\ 2 \text{ for } m > 0, l = 0 \text{ or } m = 0, l > 0 \\ 4 \text{ for } m > 0, l > 0 \end{cases}$
and the series is truncated at $K_1 - 1$ and $K_2 - 1$. If using complex basis $\phi = \left\{1, e^{i(mu+lv)}, \cdots\right\}$ $(m, l = 1, 2, \cdots)$, the surface function 2.9 can be represented as:

$$x(u,v) = \sum_{m=-(K_1-1)}^{K_1-1} \sum_{l=-(K_2-1)}^{K_2-1} g^x_{m,l} e^{i(mu+lv)}$$

$$y(u,v) = \sum_{m=-(K_1-1)}^{K_1-1} \sum_{l=-(K_2-1)}^{K_2-1} g^y_{m,l} e^{i(mu+lv)}$$

$$z(u,v) = \sum_{m=-(K_1-1)}^{K_1-1} \sum_{l=-(K_2-1)}^{K_2-1} g^z_{m,l} e^{i(mu+lv)} \tag{2.12}$$

where $g^x_{m,l}$, $g^y_{m,l}$ and $g^z_{m,l}$ are the coefficients of the corresponding harmonic indexed by m and l. Re-arrange them into a vector form:

$$g = [g^x_{m,l} \; g^y_{m,l} \; g^z_{m,l}] \; (m, l = 0, \pm 1, \pm 2, \cdots)$$

g is then a shape descriptor. The remaining part of the work in 3D is quite similar with the application in 2D. That is, the boundary finding problem is transformed into a MAP (maximum a posterior) problem. And then by searching the parameter space of g, the maximum is found. The corresponding g at which the objective function is maximized will be the boundary we are searching for.

2.3.4 SPHARM

SPHARM stands for spherical harmonics shape descriptor [41]. In this model, the surface is represented as a weighted sum of spherical harmonics which are orthogonal over the sphere: $Y_l^m(\theta, \phi)$, $-l \leq m \leq l$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$. Specifically, the basis functions are defined as [42]:

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi} \tag{2.13}$$

$$Y_l^{-m}(\theta, \phi) = (-1)^m Y_l^{m*}(\theta, \phi) \tag{2.14}$$

where $Y_l^{m*}(\theta, \phi)$ denotes the complex conjugate of $Y_l^m(\theta, \phi)$ and P_l^m denotes the associated Legendre polynomials:

$$P_l^m(w) = \frac{(-1)^m}{2^l l!} (1 - w^2)^{\frac{m}{2}} \frac{d^{m+l}}{dw^{m+l}} (w^2 - 1)^l \tag{2.15}$$

Table 2.1 gives the first few spherical harmonics up to degree 3. Fig. 2.3 shows the absolute value of the real parts of spherical harmonic basis functions up to degree 3.

In this model, in order to extend spherical harmonics to describe more general shapes, the points on the surface are firstly mapped to the unit sphere [41, 44]. The surface in $\mathbb{R}^3$ are then represented by three coordinate functions defined on the unit sphere: $(x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))$. This appropriate mapping [41, 44] is determined by iteratively solving a constrained optimization problem based on the diffusion equation. Next, to express a surface using spherical harmonics, the

Table 2.1: Analytic expressions of the first few spherical harmonics up to degree 3, in both polar (I) and Cartesian (II) form. x, y, z, and r are related to θ, and ϕ, through the usual spherical-to-Cartesian coordinate transformation. Part III of the table gives the common normalizing constants, e.g. $Y_1^0 = \sqrt{3/4\pi}\cos\theta = \sqrt{3/4\pi}z/r$

	l	m=0	m=1	m=2	m=3
I	0	1			
	1	$\cos\theta$	$e^{i\phi}\sin\theta$		
	2	$-1+3\cos^2\theta$	$e^{i\phi}\cos\theta\sin\theta$	$e^{2i\phi}\sin^2\theta$	
	3	$-3\cos\theta+5\cos^3\theta$	$e^{i\phi}(1-5\cos^2\theta)\sin\theta$	$e^{2i\phi}\cos\theta\sin^2\theta$	$e^{3i\phi}\sin^3\theta$
II	0	1			
	1	z/r	$(x+iy)/r$		
	2	$(-1+3z^2)/r^2$	$(x+iy)z/r^2$	$(x+iy)^2/r^2$	
	3	$(-3z+5z^3)/r^3$	$(x+iy)(1-5z^2)/r^3$	$(x+iy)^2z/r^3$	$(x+iy)^3/r^3$
III	0	$1/\sqrt{4\pi}$			
	1	$\sqrt{3/4\pi}$	$-\sqrt{3/8\pi}$		
	2	$\sqrt{5/16\pi}$	$-\sqrt{15/8\pi}$	$\sqrt{15/32\pi}$	
	3	$\sqrt{7/16\pi}$	$-\sqrt{21/64\pi}$	$\sqrt{105/32\pi}$	$-\sqrt{35/64\pi}$

three coordinate functions are decomposed:

$$\begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} = \begin{pmatrix} \sum\limits_{l=0}^{\infty}\sum\limits_{m=-l}^{l} c_{xl}^m Y_l^m(\theta,\phi) \\ \sum\limits_{l=0}^{\infty}\sum\limits_{m=-l}^{l} c_{yl}^m Y_l^m(\theta,\phi) \\ \sum\limits_{l=0}^{\infty}\sum\limits_{m=-l}^{l} c_{zl}^m Y_l^m(\theta,\phi) \end{pmatrix} \tag{2.16}$$

Let $c_l^m = (c_{xl}^m, c_{yl}^m, c_{zl}^m)^\top$. Then, the SPHARM expansion of the surface is:

$$\mathbf{V}(x(\theta,\phi), y(\theta,\phi), z(\theta,\phi)) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_l^m Y_l^m(\theta,\phi) \tag{2.17}$$

This representation can then be used in deformable boundary finding (model-guided segmentation) [4, 5] in the similar way as discussed in the Fourier model in Section 2.3.3.

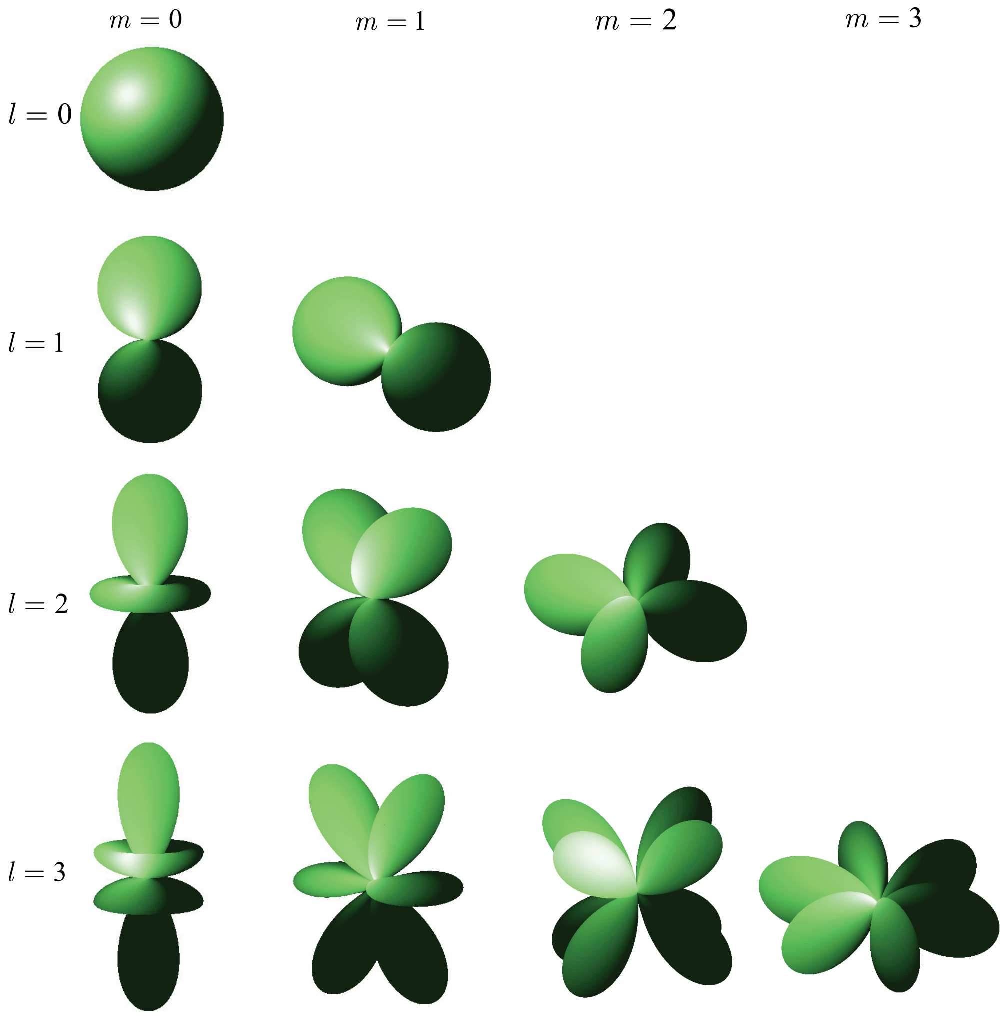


Fig. 2.3: Absolute value of the real parts of spherical harmonic basis functions up to degree 3. The figures are generated by SPHARM Generator [43]

.

2.3.5 Wavelets Based Model in 2D

Another possible set of basis functions for modal decomposition parametric model is wavelets [45, 46]. In contrast to Fourier basis functions, wavelets have compact support both in the frequency (scale) and in the spatial domain. The wavelet transform of the input signal is computed using a filter bank that splits a signal into subsampled low pass and high pass bands. This procedure is iteratively repeated for the output of low pass band. The classical wavelets are obtained by dilating and translating a fixed function, the mother wavelet. Using discrete wavelet transform (DWT) as an example, the DWT is a basis transform between certain spaces spanned by dilated and translated versions of a wavelet ψ and a scaling function ϕ:

$$\psi_i^j(x) = \psi(2^j x - i) \text{ and } \phi_i^j(x) = \phi(2^j x - i) \tag{2.18}$$

A function f is initially represented in a basis of scaling functions at a high level of resolution, denoted by the index j_ε:

$$f(x) = \sum_{i \in \mathbb{Z}} s_i^{j_\varepsilon} \phi_i^{j_\varepsilon}(x)$$

Next, a basis transform decomposes this representation into a high-frequency part, based on wavelets, and a low frequency part, based on coarser scaling functions at level $j_\varepsilon - 1$:

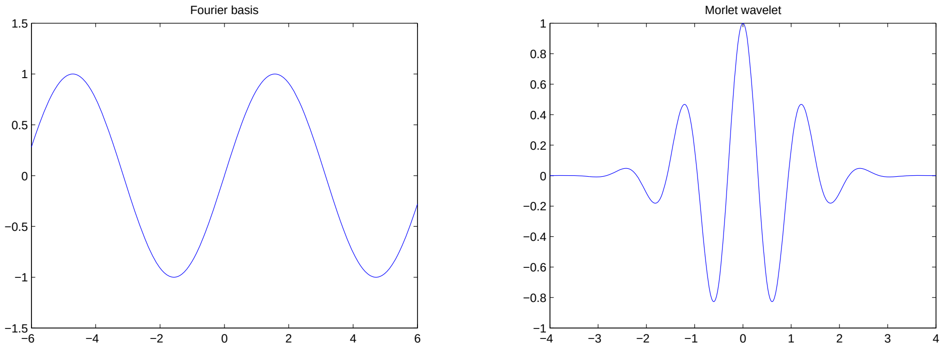
$$f(x) = \sum_{i \in \mathbb{Z}} w_i^{j_\varepsilon - 1} \psi_i^{j_\varepsilon - 1}(x) + \sum_{i \in \mathbb{Z}} s_i^{j_\varepsilon - 1} \phi_i^{j_\varepsilon - 1}(x)$$

This decomposition steps are recursively applied to the part represented by scaling functions until a base level $j = 0$ is reached. The function f is finally represented as:

$$f(x) = \sum_{j=0}^{j_\epsilon - 1} \sum_{i \in \mathbb{Z}} w_i^{j-1} \psi_i^{j-1}(x) + \sum_{i \in \mathbb{Z}} s_i^0 \phi_i^0(x)$$

By replacing the Fourier basis functions with wavelet basis functions, Chang *et al.* [24] and Davatzikos *et al.* [11] proposed wavelet descriptor for 2D boundary finding. In their work,the object boundary was parameterized and represented as 2 parametric coordinates functions, $S(x(t), y(t))$. By applying DWT decomposition, these 2 parametric functions were decomposed:

$$S(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \sum\limits_{j=0}^{N-1} \sum\limits_{i \in \mathbb{Z}_j} a_i^j \psi_i^j(t) + \sum\limits_{i \in \mathbb{Z}_j} b_i^0 \phi_i^0(t) \\ \sum\limits_{j=0}^{N-1} \sum\limits_{i \in \mathbb{Z}_j} c_i^j \psi_i^j(t) + \sum\limits_{i \in \mathbb{Z}_j} d_i^0 \phi_i^0(t) \end{bmatrix} \tag{2.19}$$

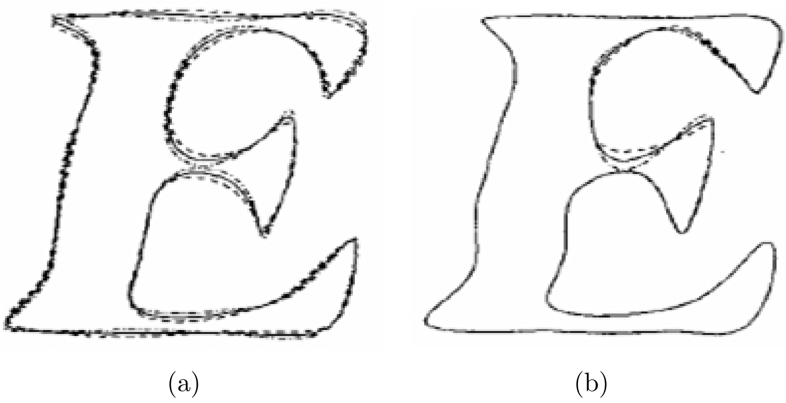


(a) (b)

Fig. 2.4: Fourier basis function vs. Wavelet basis function. (a) the Fourier basis function. (b) the Morlet wavelet

where N is the number of decomposition scales and $\mathbb{Z}_j = \{0, 1, \cdots, 2^{|j|} - 1\}$. Using this new shape descriptor, the boundary finding problem was formed into a MAP (maximum a *posterior*) problem, and convincing results were achieved.

Moreover, because the wavelet descriptor uses a set of basis functions with local support (Fig. 2.4(b) shows one example of Morlet wavelet [47].). Therefore, the shape descriptor based on wavelet decomposition provides a scheme to model local as well as global deformation. To demonstrate this ability, after changing the value of a certain coefficient of the wavelet descriptor, the object boundary is reconstructed and shown in Fig. 2.5(b), in which we can see that only part of the boundary is affected. However, because the basis functions of the Fourier descriptor are the sinusoids which are periodic and global supported (not sufficiently localized in space as shown in Fig. 2.4(a)), so that a change of one coefficient will affect the entire outline of the reconstructed boundary as shown in Fig. 2.5(a). Therefore, Fourier deformable model is not efficient in describing shapes with only local deformation. By comparing Fig. 2.5(a) and Fig. 2.5(b), the superior local deformation property of the wavelet descriptor can be easily seen.



(a) (b)

Fig. 2.5: Shape descriptors: globally supported vs. compactly supported. (a) boundary reconstructed from Fourier descriptor. (b) boundary reconstructed from wavelet descriptor. (adopted from [24])

2.4 Comparison Between Different Models

After introducing the related work in shape representation, it is worthwhile to compare these existing statistical/deformable shape models, before our new model is proposed.

2.4.1 The Selected Properties of a Shape Model

To make the comparison clearer, firstly, a selection of general shape description properties that can be used to evaluate most shape description methods are listed:

A. Multiscale shape representation

When the anatomical objects are segmented by human expert's interaction delineation, because of the presence of noise, partial volume effects, intensity inhomogeneities and other artifacts, the manual segmentation can't be considered as error-free, especially in the small scales. Therefore, a too detailed description is in fact not appropriate, since it reconstructs the objects to an unnecessarily high degree of precision based on the non-accurate manual segmentation. On the other hand, it is also desirable to be able to precisely detect shape changes, so that a representation of anatomical structures in details is also needed. Thus, the choice of scale can be interpreted as balancing the trade-off between the efficiency of description and accuracy when locating the shape variance. Moreover, if the shape descriptor can represent the shape in a hierarchical way, this feature will result in a hierarchical segmentation process in model-guided segmentation.

B. Efficient shape description

The efficiency of a shape description is defined as the amount of information required to describe the shape at certain specified description accuracy. A shape description is called efficient if objects are described with a given accuracy by concise sets of parameters or features. Implicit or explicit list of points or mesh facets will only result in a verbose description of the shape. Therefore, a more compact representation is desired to reduce the dimension of the shape model.

C. Spatial localization

This property refers to whether a representation captures the object as a set of coefficients of basis functions with locality. If a description is lacking of this locality, the local shape deformations will not result in changes of only part of coefficient set but rather in changes over the whole set of coefficients. Thus, changes of the coefficients cannot be interpreted intuitively. Therefore, for the parametric models, the support of the basis function is better to be localized.

D. Continuous shape description

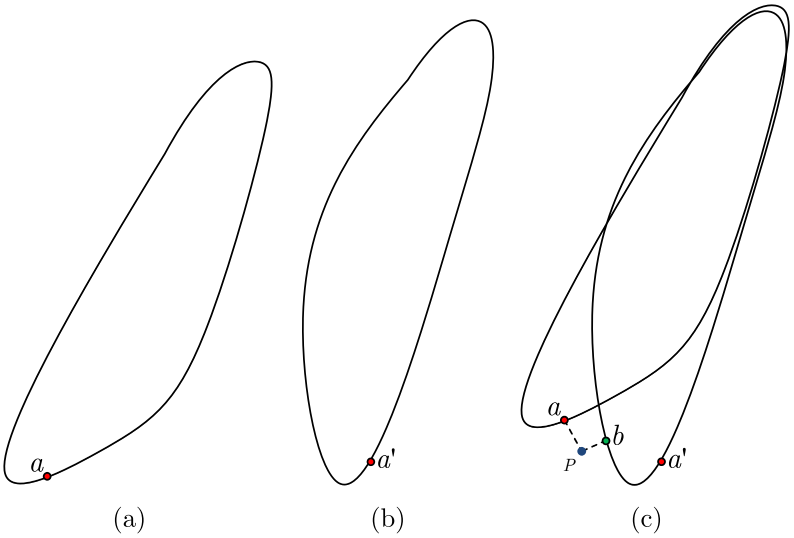
The shape should be defined continuously other than only known at discrete locations.

E. Adaptive topology

This feature refers to the ability of the shape model to change topology adaptively in the deformation process. Although the topologies of most of the biological structures are consistent from person to person, there do exist some structures, for example vasculature, have different topologies on different persons.

F. Easy to establish appropriate correspondence

Correspondence which defines the homology of points between different objects, is extremely important in order to compare shapes and generate statistics.



(a) (b) (c)

Fig. 2.6: Problematic correspondence. (a) boundary of object 1; (b) boundary of object 2; a and a' are two corresponding points on these two boundaries. (c) the result of Procrustes alignment. P is a point on the distance map where the value of the distance transform is determined as the distance $-|aP|$ and $-|bP|$ respectively in 2 distance maps. So, when the average distance map at P is computed, point a and b are, in fact, treated as two corresponding points.

For the shape models which represent the object boundary directly, the correspondence is usually established by finding the corresponding anatomical landmarks on different objects either by manual labeling or by some automatic computation. However, for some shape models which are not based on the direct description of object boundary, the defining of correspondence may be problematic. For example, in the distance transform/level set model, it is quite hard to establish appropriate correspondence between the distance maps. We use a 2D case as an example to illustrate. Fig. 2.6(a) and Fig. 2.6(b) shows the boundaries of two objects in the training set. a and a' are the two corresponding points on these two boundaries. However, after these two boundaries are Procrustes aligned (Fig. 2.6(c)), at point P, the corresponding point of a on the distance map is actually defined as point b. Therefore, because of this reason, the correct mean shape can not be acquired by averaging the distance maps.

Table 2.2: Comparison between different shape models. A-F represent the 6 most desirable properties discussed in Section 2.4.1: A. Multiscale shape representation; B. Efficient shape description; C. Spatial localization; D. Continuous shape description; E. Adaptive topology; F. Easy to establish appropriate correspondence. ’√’ denotes that the model has the desirable property, otherwise ’×’ denotes not.

	A	B	C	D	E	F
PDM	×	×	√	×	×	√
Level Set	×	×	√	×	√	×
Discrete Mesh	×	×	√	×	×	√
ASM	×	√	×	×	×	√
SPHARM	√	√	×	√	×	√
Fourier descriptor	√	√	×	√	×	√
Wavelet descriptor	√	√	√	√	×	√

2.4.2 Comparison Between Different Shape Descriptions

Having discussed the desirable properties of shape representations used for building statistical shape model in the previous section, different shape models can be compared based on these properties. The comparison results are summarized in Table 2.2.

From Table 2.2, we can see that the parametric models are superb than free-form models in most of the comparisons. The parametric models usually provide efficient, multiscale and continuous shape representations which can not be achieved by free-form models. The level set model is one special free-form model, in which the shape is not represented explicitly but implicitly embedded as a zero level set of the distance transformation. Therefore, it has the ability to change the topology adaptively during the deformation and is quite useful to segment objects, such as vasculature, with variable topology. However, because the topology of most biological objects is only a sphere [5], this property is not necessary for the analysis of most biological objects.

The Active Shape Model (ASM) [2] is a special case of parametric models, because although its shape representation is in a parametric form (a linear combination of eigenvectors of training set), it still uses an explicit list of points on the boundary to describe the shape. Thus, we know the geometry of a shape only at a finite set of points. When a high degree of precision is required on the shape, a corresponding dense sampling is needed. This shape model can be verbose and thus not efficient for computation. Moreover, as the training set is usually small in size, relative to the number of sample points on the surface, the possible ways to deform such a shape model are limited to a linear subspace of the complete shape space.

The comparison between wavelet-based model and other parametric models shows that, besides the common advantages of parametric models (multi-scale, efficient and continuous) the wavelet model gives the extra valuable spatial-localization property because of its compactly supported basis functions. The spatial-localization property can lead to the important ability of modeling the locale shape variations and performing a spatially localized model-image fitting.

2.5 Extend the Wavelet Model to 3D

Compared with Fourier or spherical harmonics, as we can see, wavelet basis has many advantages both in spatially localized shape variation representation and the later model-image fitting process. Hence, a wavelet shape model in 3D for model-guided segmentation and statistical shape analysis will be highly desirable. However, to extend the existing wavelet model from 2D boundary to 3D surface is severely hampered, because the classical wavelet used in [11, 24] is a scaled and translated copy of a single unique function (mother wavelet), but, this shift-invariant theory breaks down when representing functions defined on a bounded surface. Another barrier preventing the extending of classical wavelets scheme to 3D surface is the rigorous requirements in the explicit parametrization on 3D bounded surfaces. We need, firstly, a surface parametrization that provides the correspondences between objects in the training set. Secondly, the first-generation wavelets require the surface to be parameterized by regular grids. Since the topology of anatomical structures is usually a sphere, we need two parameters, longitude and latitude to characterize such a surface. However, in doing so, distortions are inevitably generated at the south and north poles.

To address this issue, we propose a novel Statistical Surface Wavelet Model (SSWM) and a corresponding model-guided segmentation method based on a recently introduced wavelet decomposition and synthesis scheme [48] defined directly on the subdivision surface mesh. It has many desirable properties for building a shape model, performing model-guided segmentation and statistical shape analysis. Details of this wavelet scheme and how to use it for statistical shape model building is discussed in next chapter.

2.6 Recent Related Work

Near the end stage of this thesis work, we also found two closely related works were published very recently. Both are based on the spherical wavelets [49], another subdivision surface wavelet scheme defined exclusively on the unit sphere. Using the spherical wavelets, in [50, 51], the author proposed a multiscale and spatially localized analysis method of the cortical surfaces reconstructed from the MR images of a set of subjects, which is impossible for the other global basis set, such as spherical harmonics. The work in [52–54] ismore related to ours. They also developed a model-guided segmentation framework to utilized the "double localization" property of wavelets. So it is worthwhile to compare these two methods. In general, there are 3 main differences. Firstly, our wavelets are defined on a different subdivision scheme based on the Catmull-Clark subdivision [48] which is capable of representing any two-manifold geometries as well as functions defined on these, whereas the mentioned work is on a triangulated subdivision wavelets defined on the unit sphere. Secondly, our model contains the complete biological variation, inclusive of the shape variations and variations in similarity transform (translation, rotation and scaling). As what will be mentioned in Section 3.2.3, such a prior model integrating similarity transform is much more robust than a pure "shape" model [4, 5]. This is because a pure "shape" model loses the correlation between the position and the shape of the anatomical structure. Therefore, with our model, unlike that in the mentioned work, we do not need the additional step of a pre-optimization of the pose parameters to initialize the model. Lastly, the mentioned work employed a local search gradient descent algorithm. In contrast, we capitalize fully on both the frequency and spatial localization properties to develop a simple and efficient equal sampling algorithm which is robust, in particular, for cases where there are local optimums near to initial positions.

Chapter 3

Statistical Surface Wavelets Model (SSWM)

In this chapter, we introduce a novel statistical shape model—The Statistical Surface Wavelet Model (SSWM). The main purpose of this model is to extend the wavelet-based statistical shape modeling method to 3D surfaces using the second generation wavelets defined on the subdivision surface mesh, which has many advantages over the other methods. In this chapter, we will first introduce the surface wavelet decomposition scheme we adopted. After that, the other two pivotal issues in the statistical shape model building, i.e. correspondence finding and re-meshing will be discussed. Lastly, a SSWM of the caudate nucleus used for model-guided segmentation is built based on a training set using the proposed method.

3.1 The Shape Representation Based on Subdivision Surface Wavelets

In this section, we will first introduce the subdivision surface wavelet scheme we adopted in our proposed statistical shape model, Generalized B-Spline Subdivision-Surface Wavelets [48]. Next, as shape representation in statistical shape model building, the new shape representation based on subdivision surface wavelets will be compared with other shape representation methods with different basis functions.

3.1.1 The Related Work

Wavelet, because of its many advantages in level of details analysis, has been applied in many domains since 1980s. However, there was no surface wavelet scheme for computer graphics reported until a later time. Wavelets representing surfaces of arbitrary topology were originally explored by Lounsbery *et al.* [55, 56]. Starting with a subdivision surface scheme, like Catmull-Clark [57] or Loop [58] subdivision, wavelet transforms were constructed using the recursively generated basis functions as scaling functions. However, these wavelet constructions have the disadvantage that the transform is based on a global system of equations. Only the inverse transform can be generally computed in linear time (in terms of number of vertices) based on local operations. On the other hand, as we know, the model-guided boundary finding problem is often formulated as an optimization problem. The wavelets decomposition process and synthesis process are repeated many times in the optimization stage. Therefore, this kind of application is sensitive to the computational cost of the surface wavelets decomposition and synthesis. Hence, this method is not suitable for the shape representation purpose in building a statistical shape model and used for model-guided segmentation.

Later, another subdivision-surface wavelet constructed for functions defined on triangulated spherical domains were introduced by Schröder and Sweldens [49]. In this method, subdivisions and wavelets are smoothly joined together by the *Lifting Scheme* [25], an efficient technique for constructing second generation wavelets. Schröder and Sweldens [24] showed an example of constructing lifting wavelets on the sphere. This method has linear time decomposition and synthesis. However, their approach is used for constructing functions on the given sphere domain rather than representing the underlying domain geometries (other than sphere).

Until recently, Bertram *et al.* [48] introduced a new biorthogonal wavelet decomposition and synthesis scheme based on generalized B-subdivision surfaces. This method also uses *Lifting Scheme* [25], but is capable of representing two-manifold geometries as well as functions defined on the two-manifold. More importantly, both its decomposition and synthesis are linear to the number of points on the mesh.

3.1.2 The Generalized B-spline Subdivision-Surface Wavelets

The subdivision-surface wavelet we used [48] is a second-generation wavelets scheme. It is based on the hierarchical mesh connectivity defined by Catmull-Clark subdivision [57]. Fig. 3.1 shows an example, where we call the initial mesh *sub-mesh*(right) and the mesh resulting from subdivision *super-mesh*(left), since it contains more vertices. A mesh is refined (from sub-mesh to super-mesh) by inserting a new vertex inside every face and on every edge and by connecting these vertices to quadrilaterals. Vertices in a super-mesh correspond to a face (polygon), an edge, or a vertex in the sub-mesh and are denoted by $\mathbf{f}$, $\mathbf{e}$, and $\mathbf{v}$, respectively. A decomposition step can be considered as an operation applied to a super-mesh that computes $\mathbf{v}'$ vertex positions for an approximating sub-mesh and replaces the remaining $\mathbf{e}$ and $\mathbf{f}$ vertices by difference vectors ($\mathbf{e}'$ and $\mathbf{f}'$) which represent details that are missing in the sub-mesh. The $\mathbf{v}'$ vertices represent coefficients for scaling functions, $\mathbf{e}'$ vertices represent wavelet coefficients corresponding to edges, and $\mathbf{f}'$ vertices represent wavelet coefficients corresponding to faces. The reconstruction (synthesis) is then a step of reconstructing the super-mesh from the sub-mesh by adding the detail shape information contained in the wavelet coefficients $\mathbf{e}'$ and $\mathbf{f}'$.

If we denote the set of all vertices contained in the mesh after j subdivisions as $\mathbf{V}(j)$. When we obtain a finer resolution mesh $\mathbf{V}(j+1)$ through subdivision, we denote the $\mathbf{e}$ vertices and $\mathbf{f}$ vertices added to the subdivision as the vertex set

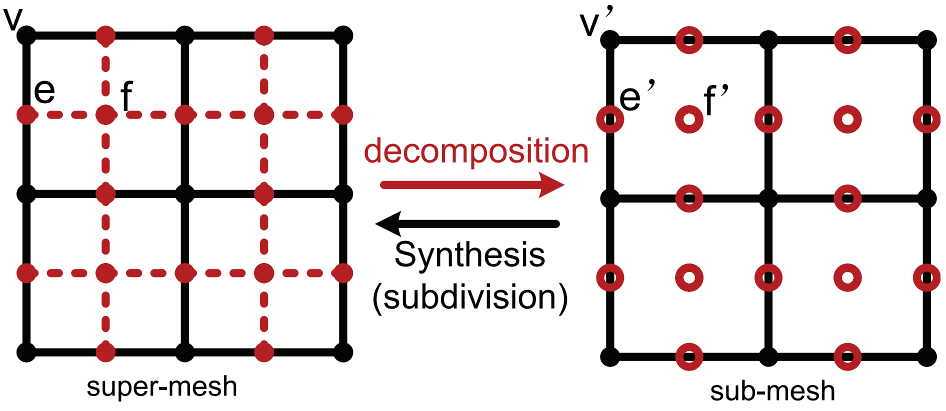


Fig. 3.1: Wavelet transformation on Catmull-Clark subdivision mesh. $\mathbf{f}$, $\mathbf{e}$, and $\mathbf{v}$ represent the vertices in a super-mesh which correspond to a face (polygon), an edge, or a vertex in the sub-mesh, respectively. $\mathbf{v}'$ is the vertex on the approximating sub-mesh. $\mathbf{e}'$ and $\mathbf{f}'$ are the difference vectors which represent details that are missing in the sub-mesh.

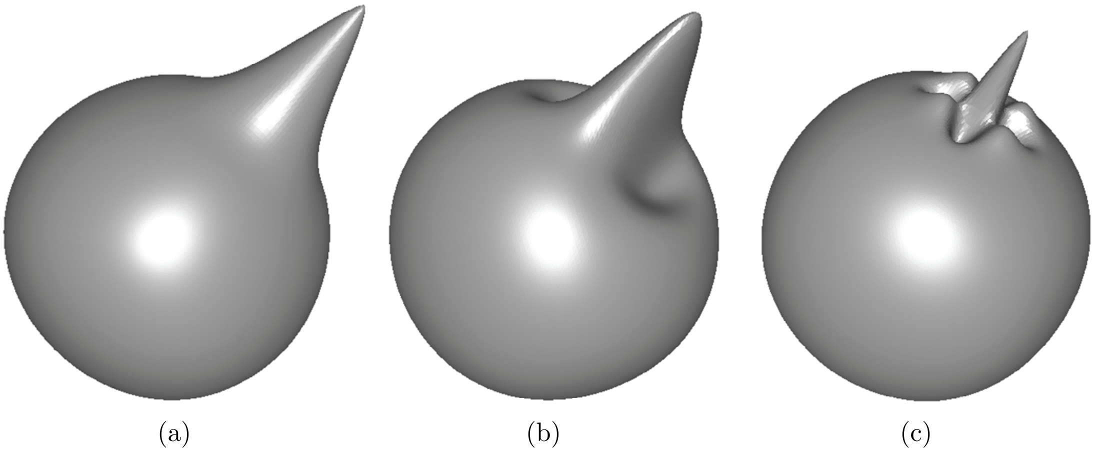


Fig. 3.2: Basis functions on a sphere. (a) Scaling function. (b) Wavelet function corresponding to an edge. (c) Wavelet function corresponding to a face.

$\mathbf{W}(j)$. The complete set of vertices in the finer resolution $j+1$ is $\mathbf{V}(j+1) = \mathbf{V}(j) \cup \mathbf{W}(j)$. Let S be a surface and $\mathbf{x}(x, y, z) \in \mathbb{R}^3$ a vertex on S. Using wavelet transform described above, S can be represented by a weighted summation of a set of scaling and wavelet functions (shown in Fig. 3.2) defined on the surface with different scales and locations:

$$S(\mathbf{x}) = \sum_{j \geq 0} \sum_{m \in \mathbf{W}(j)} \mathbf{w}_m^j \cdot \Psi_m^j(\mathbf{x}) + \sum_{n \in \mathbf{V}(0)} \mathbf{v}_n^0 \cdot \Phi_n^0(\mathbf{x}) \tag{3.1}$$

where, Φ_n^0 is the scaling function of the coarsest scale at vertex n, Ψ_m^j is the wavelet function of scale j at vertex m, and $\mathbf{v}_n^0$ and $\mathbf{w}_m^j$ are the corresponding coefficient vectors of these basis functions. In 3D, $\mathbf{v}_n^0$ and $\mathbf{w}_m^j$ are column vectors with 3 elements and each element represents one of the coordinates x, y, z. For detail of the transformation please refer to Appendix A.

3.1.3 Surface Wavelets as Shape Descriptor

The set consisting of all the wavelets coefficient vectors in (3.1):

$$C = \{\mathbf{v}_n^0, \mathbf{w}_m^j | j = 0, 1, 2, \cdots; n \in \mathbf{V}(0); m \in \mathbf{W}(j)\} \tag{3.2}$$

is, in fact, a shape descriptor, because from it the shape can be reconstructed exactly.

As mentioned in Section 2.4, the ability to represent the object surface at different scale levels is quite crucial for a shape model used in medical image analysis. To get the shape representation at different resolution is very simple using the subdivision surface wavelet shape model—just omitting the wavelet coefficients in the unwanted scale levels. Fig. 3.3 shows an example of a multiresolution shape representation of cerebral lateral ventricle. As we can see, from Fig. 3.3(b) to Fig. 3.3(f), more and more detailed shape information is added to the representation.

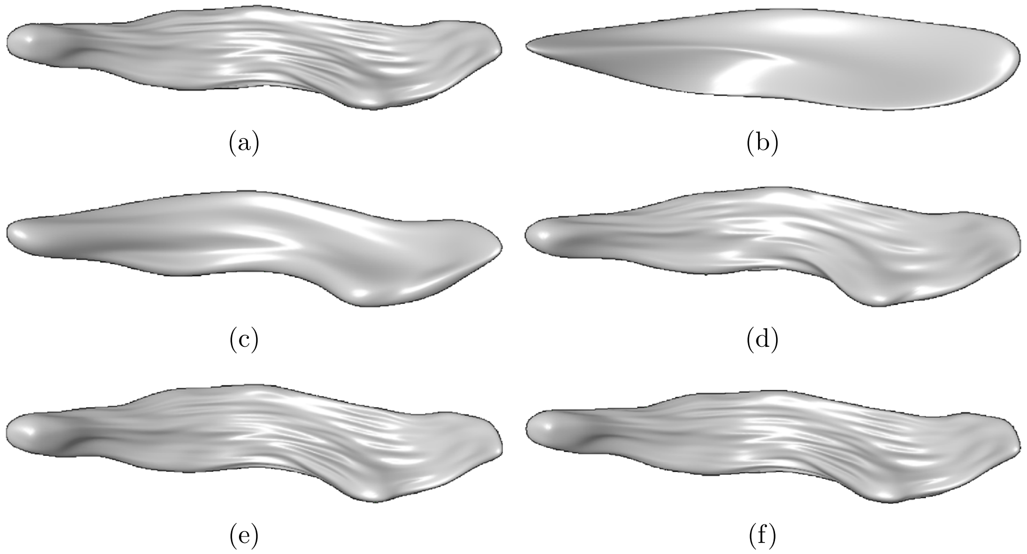
Beside the multiscale shape representation, the wavelet based shape model provides another superb spatial-localization property because of the compactly supported basis functions (shown in Fig. 3.2). To demonstrate this advantage clearly, we can also use the cerebral lateral ventricle as an example. In Fig. 3.4(a), the original surface is shown. Next, this surface is decomposed and represented by a set of spherical harmonics coefficients. After that, a simulated shape variation is generated by changing the spherical harmonics descriptor coefficient of degree 17 order 1 and the result is shown in Fig. 3.4(b). As we can see, because the globally supported basis function of spherical harmonics (shown in Fig. 2.3), the change of only one coefficient has changed the entire outline of the shape. Since this shape representation does not provide spatially localized shape information, there are 2 disadvantages when it is used in building a statistical shape model. Firstly, as we know that some diseases, such as cancer, may only affect the shape of some part of an organ. With these models without spatial-localization, from the variation of coefficients we could not know the location of the shape variations. Secondly, since the coefficients in the same scale level are correlated, all the coefficients in one scale must be optimized together in the model-guided segmentation process.

However, in contrast to spherical harmonics shape descriptor, the change of one coefficient in the wavelet shape descriptor only change the shape within the scale level and location where the shape is defined by the chosen coefficient, whereas the remaining part of the surface keeps unchanged (as shown in Fig. 3.4(c)). Inversely, if any shape variations are detected by comparing the wavelet coefficients, the scale and location information about the shape variation can be deduced directly from the wavelets coefficients.

Since the shape representation of surface wavelets has so many advantages, we will propose a novel statistical shape model based on this shape representation in the next sections. But before that, other 2 key problems in building a statistical shape model, correspondence finding and re-meshing, have to be discussed and solved.

3.2 The Correspondence Finding and Re-meshing Problem

Correspondence finding and re-meshing are two common and very important problems in building a statistical shape model. Fig. 3.5 shows an example training set used for the construction of a statistical shape model of the caudate nucleus. As we can see, the binary volumetric images consist of different number of vox-

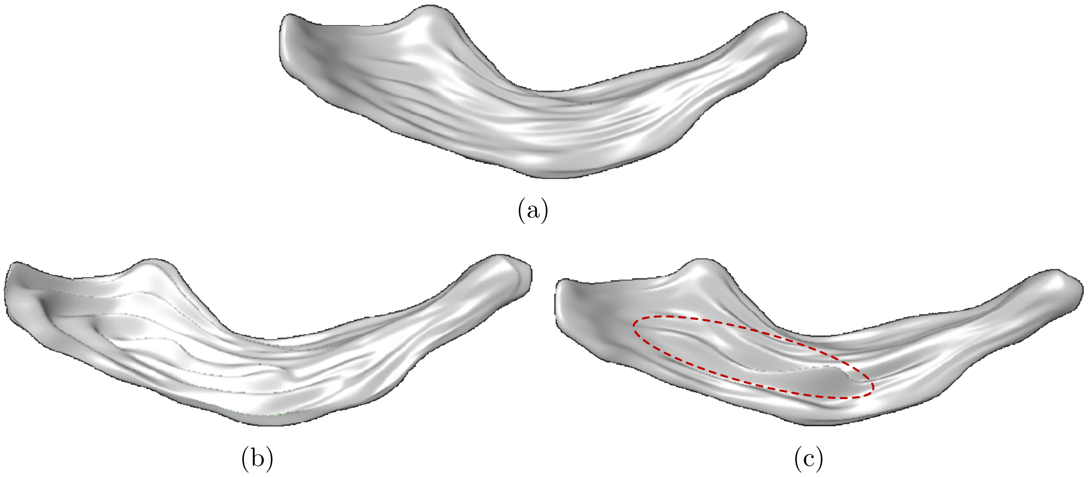


(a) (b)

(c) (d)

(e) (f)

Fig. 3.3: Multiscale representation of the cerebral lateral ventricle using the subdivision surface wavelets: (a) the original shape; (b)-(f): shape reconstruction up to scale level 0,1,2,3 and 4, respectively.



(a)

(b) (c)

Fig. 3.4: spatially localized shape representation: (a) the original shape; (b) a variation generated by changing SPHARM descriptor coefficient of degree 17 order 1; (c) a variation generated by changing a wavelet coefficient corresponding to an edge at scale level 2.

els and are in different orientation, translation and scale. Therefore, before any comparison between them can be performed, the corresponding pairs of points on different surfaces must be defined. Moreover, in our work, since the wavelet transform can only be conducted on mesh with the required mesh connectivity, a preparation step—re-meshing is also necessary. After re-meshing, the points on

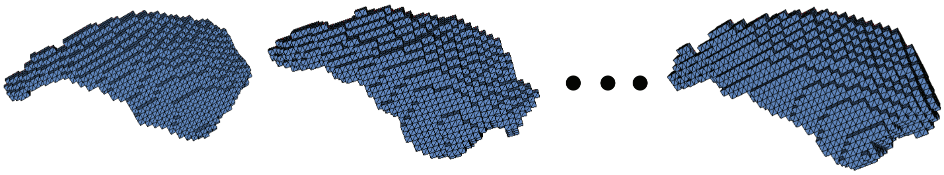


Fig. 3.5: Segmented binary volumetric data.

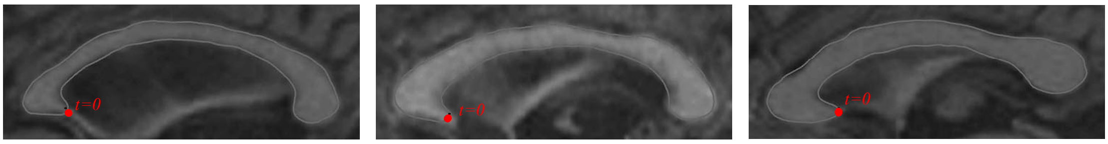


Fig. 3.6: Correspondence finding in 2D boundary through parameterization. t is the parameterization variable. The points on different object boundaries with the same t value in parameterization are defined as corresponding points.

the re-meshed surfaces with the same mesh vertex index must be the corresponding points. So, in fact, correspondence finding and re-meshing are two closely correlated problems. In this section, our method of solving these two problems in a single framework is discussed. Before that, some related work in correspondence finding will be introduced firstly.

3.2.1 Related Work

Correspondence finding, in general, is a difficult problem. While shape features provide the basis for such a correspondence, this problem remains a difficult one due to the ambiguity in finding the landmarks when the surfaces are complex and variable. Moreover, because the number of the reliably identifiable landmarks is too small to describe a shape (especially for surfaces, usually hundreds or thousands points are needed), manual landmarking becomes impractical. In addition, the lack of ground truth of the correspondence remains a problem for evaluating these methods.

A number of methods have been developed to match contours and surfaces to determine the correspondences. The classic approach is the iterative closest point

(ICP) method [59]. This method minimizes the distance from points in one surface (contour) to the closest point in another surface. The main problem in this method is that the corresponding point on another surface is probably not a mesh vertex but a point inside a mesh facet. Recently, many improved automated approaches were proposed. These approaches formulate the correspondence problem as that of defining a parameterization for each surface (contour) in the training set and assume correspondence between equivalently parameterized points (as shown in Fig. 3.6). Kotche *et al.* [60] proposed to use an optimization process to assign the best correspondence across all objects of the training population. It is based on the assumption that the correct correspondences are, by definition, those that build the optimal average model given the training population. For that purpose they proposed to use the determinant of the covariance matrix (DetCov) as an objective function. Later, based on the idea of the DetCov method, Davies proposed a different objective function [61] for the optimization process using the Minimum Description Length (MDL [62]) principle. Another approach was proposed by Wang *et al.* [63] based on manually initialized subdivision surfaces. This method is semi-automatic starting from a set of pre-defined anatomical landmarks which is triangulated in a standardized manner. The correspondences between points other than the pre-defined landmarks are defined by further refined subdivision surfaces. An evaluation of the different 3D correspondence methods for statistical model building is available in [64].

As discussed in Section 2.3.4, the Spherical Harmonic (SPHARM) description [41] is a parametric surface description using spherical harmonics as basis functions. Correspondence finding using SPHARM is a method which solves the correspondence problem not in the shape space but in the spectral space of spherical harmonics through a normalization process using the rotation-invariant property of spherical harmonics to exclude translation, rotation, and scaling. In our work, for the genus-zero object, we adopt this method to define the correspondence. The reason why we choose this method is that from the results of SPHARM normalization, a set of normalized spherical harmonics coefficients, we can acquire the required mesh connectivity directly by reconstructing the surfaces using a resampling grid on unit sphere with Catmull-Clark subdivision connectivity. Next, an example will be given to explain how this method works.

3.2.2 Correspondence Finding and Re-meshing Through SPHARM Normalization

In this section, we will use an example to illustrate the correspondence finding and re-meshing method we adopted. Because lateral ventricle has a more complex and tortuous shape, it is more difficult to get a good spherical mapping (critical part of the correspondence finding problem) onto the unit sphere. Therefore, we use lateral ventricle as the example to show the effectiveness of the correspondence finding method. Our work uses anatomical objects obtained from MR images. After binary volumetric data of the interested structure is obtained through a segmentation process, the outermost facets of the voxels are extracted as the surface mesh of the anatomical object. Fig. 3.7(a) and Fig. 3.7(b) gives two such examples of the extracted surface meshes of lateral brain ventricle. In the following, we describe the three steps in the SPHARM normalization correspondence finding method: (1) unit sphere mapping, (2) SPHARM expansion, and (3) SPHARM normalization.

Unit-sphere mapping aims to create a continuous and uniform mapping from the object surface to the surface of a unit sphere. This mapping is formulated as a constrained optimization problem with the goals of topology and area preservation and distortion minimization [41, 44] (executable of this algorithm is available from [65]). The result is a bijective mapping (shown in Fig. 3.7(c) and Fig. 3.7(d)) between each point $\mathbf{V}(x, y, z)$ on a surface and two spherical coordinates θ and ϕ on a unit sphere:

$$\mathbf{V}(x,y,z) = \begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} \tag{3.3}$$

where x, y and z denote the Cartesian object space coordinates and θ and ϕ are the two spherical coordinates. When the free variables θ and ϕ run over the whole sphere, $\mathbf{V}(x,y,z)$ runs over the whole surface of the input object.

SPHARM expansion expands the object surface into a set of SPHARM basis functions Y_l^m, where Y_l^m denotes the spherical harmonic of degree l and order m. The spherical harmonics are defined as:

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi} \tag{3.4}$$

where $P_l^m(\cos\theta)$ are associated Legendre polynomials (with argument $\cos\theta$), and l and m are integers with $-l \leq m \leq l$. After the unit sphere mapping, the Cartesian object space coordinates x, y and z of the points on the surface are functions of variables θ and ϕ defined on the unit sphere. Thus, each of them can be decomposed with the SPHARM basis functions:

$$\begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} = \begin{pmatrix} \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{xl}^m Y_l^m(\theta,\phi) \\ \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{yl}^m Y_l^m(\theta,\phi) \\ \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{zl}^m Y_l^m(\theta,\phi) \end{pmatrix} \tag{3.5}$$

Let $\mathbf{c}_l^m = (c_{xl}^m, c_{yl}^m, c_{zl}^m)^\top$. Then, the surface SPHARM expansion is:

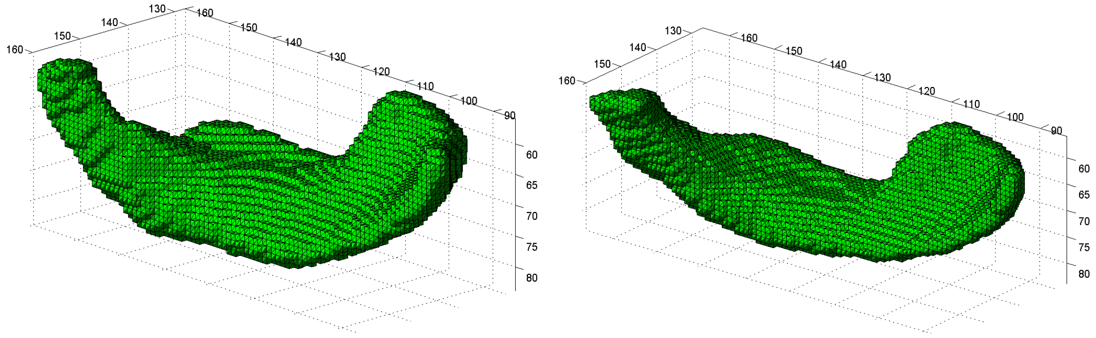
$$\mathbf{V}(x(\theta,\phi), y(\theta,\phi), z(\theta,\phi)) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \mathbf{c}_l^m Y_l^m(\theta,\phi) \tag{3.6}$$

Given these coefficients $\mathbf{c}_l^m$, by re-sampling on the unit sphere to specify θ and ϕ

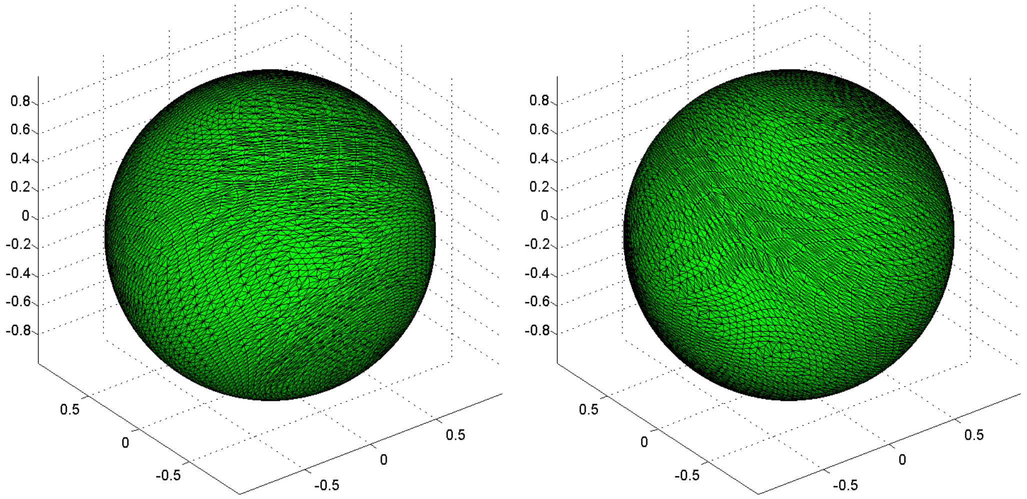
in the right side of Eq. (3.6), inversely, we can reconstruct the surface. And using more coefficients leads to a more detailed reconstruction.

SPHARM normalization is a very similar process as Fourier descriptor normalization [66]. It removes translation, rotation, and scaling in objects to generate a normalized set of SPHARM coefficients, which are comparable across different objects. In brief, rotation invariance is achieved by aligning the degree 1 reconstruction, which is always an ellipsoid. The parameter net on this ellipsoid is rotated to a canonical position such that the north pole is at one end of the longest main axis, and the crossing point of the zero meridian and the equator is at one end of the shortest main axis. In the object space, the ellipsoid is rotated to make its main axes coincide with the coordinate axes, putting the shortest axis along x and longest along z. Scaling invariance is achieved by dividing all the coefficients by a scaling factor, and translation invariance is achieved by ignoring degree 0 coefficient. After the above steps, a set of normalized coefficients $\mathbf{c}_l^{\prime m} = (c_{xl}^{\prime m}, c_{yl}^{\prime m}, c_{zl}^{\prime m})^\top$, $(-l \leq m \leq l,\ m = 0, 1, 2...)$ can be obtained to form a normalized shape descriptor for each object surface, in which corresponding surface points can be defined using parameter θ and ϕ. That is, points on two surfaces with the same spherical coordinates θ and ϕ are corresponding points.

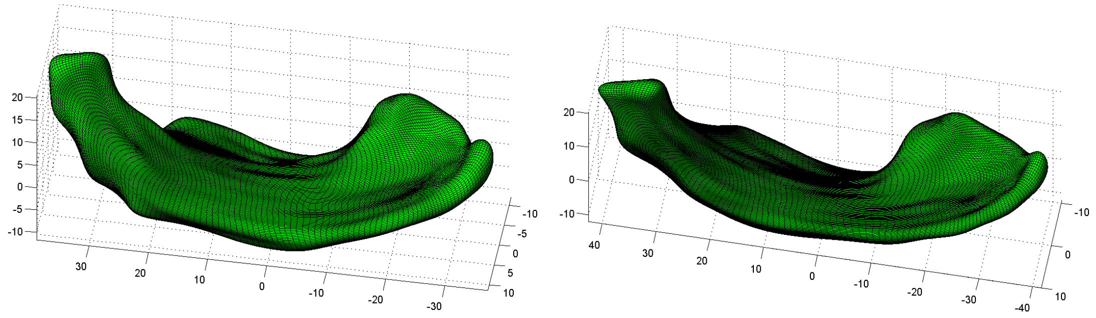
The above three steps are the standard steps in SPHARM normalization [41]. But after these steps, in order to get the required mesh connectivity, we add the re-meshing as a following step to reconstruct the surface meshes from these normalized coefficients $\mathbf{c}_l^{\prime m}$ by specifying the values of θ and ϕ in the right side of (3.6). Since in the above SPHARM normalization process, translation, rotation and scaling factors has been excluded, surface reconstruction using a standard re-sampling grid on the unit sphere with Cutmall-Clark subdivision mesh connectivity (as shown in Fig. 3.8) will result in reconstructed surfaces with the same mesh connectivity. And in the reconstructed surfaces, the points on different meshes with the same mesh index are the corresponding points because they have the same values of θ and ϕ. Fig. 3.7(e) and Fig. 3.7(f) shows the normalized and re-meshed



(a) cerebral lateral ventricle 1 (b) cerebral lateral ventricle 2



(c) unit sphere mapping of ventricle 1 (d) unit sphere mapping of ventricle 2



(e) normalized and re-meshed surface of ventricle 1 (f) normalized and re-meshed surface of ventricle 2

Fig. 3.7: The SPHARM normalization and re-meshing.

surfaces of the two cerebral lateral ventricles shown in Fig. 3.7(a) and Fig. 3.7(b), respectively.

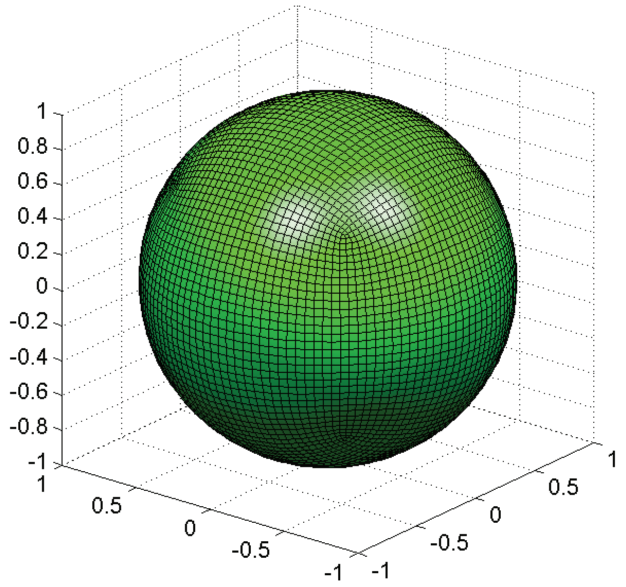


Fig. 3.8: The sampling grid with Catmull-Clark subdivision mesh connectivity.

3.2.3 Talairach Coordinates and a Shape Prior Integrating Similarity Transform Information

After the steps described in last section, the surface mesh has been re-meshed and the correspondence between different surfaces has been established as well. As mentioned, for each shape, the SPHARM normalization removes its translation, rotation and scaling, or collectively its similarity transform. As such, it also removes the correlation between the similarity information and the shape of the anatomical structure (in the original images). This effect is clearly visible through comparing the coordinates labels of the bounding box in Fig. 3.7(a) and Fig. 3.7(b) with that in Fig. 3.7(e) and Fig. 3.7(f), respectively. Note that the surface mesh has been moved so that the center of mass coincide with the coordinate origin. The orientation of the surface is also rotated to coincide with the direction of y axis. However, it was observed in [5] [4] that a prior model integrating similarity transform is much more robust than a pure "shape" model in model-guided segmentation. Therefore, in our framework, we use a simple principal component alignment method to register a surface obtained from the SPHARM normalization with its original surface obtained directly from the input binary segmented image to restore its similarity transform information in the standard stereotactic coordinate system.

Here, this standard stereotactic coordinate system we adopted is the Talairach coordinate system [67]. It was proposed by Talairach and Tournoux for global alignment of the head image data sets. Basic landmarks used to establish Talairach coordinate system are the inter-hemispheric commissures in the mid-sagittal plane, the so-called anterior and posterior commissures (AC/PC) (see Fig. 3.9(a)). As shown in Fig. 3.9(a)-Fig. 3.9(d), in this coordinate system, the origin is set at the AC point. Axis y traverses the brain through AC-PC line from back (negative values) to front crossing the two feature points. Also in the mid-sagittal plane, axis z is perpendicular to y and points from bottom to top. The third axis x, pointing from left to right, is defined to be orthogonal to the other two such that they build a right handed coordinate system. For each data set, once Talairach coordinate system is established, it is transformed into this canonical coordinates by 3D rotation and scaling.

Therefore, to transform the images in the training set into the T-T coordinate, we need to determinate the mid-sagittal plane and the position of the AC/PC line in every training samples. Currently this process is done manually. To derive the position of the mid-sagittal plane, firstly, the position of three points lying on the inter-hemispheric plane are specified. Then, the plane defined by the these three points—the mid-sagittal plane is computed. Having determined the plane, then, locations of the AC and PC are marked. Lastly, a translation vector and a rotation matrix are computed to transform the coordinate system in every training sample into the Talairach system. In the future, the determination of the symmetry mid-sagittal plane and the AC/PC line can be replaced by automatic methods [68–72].

After defining the Talairach coordinates system, the next step is to register the SPHARM normalized surface mesh (as shown in Fig. 3.7(e) and Fig. 3.7(f)) with the corresponding original surface of binary volumetric image (as shown in

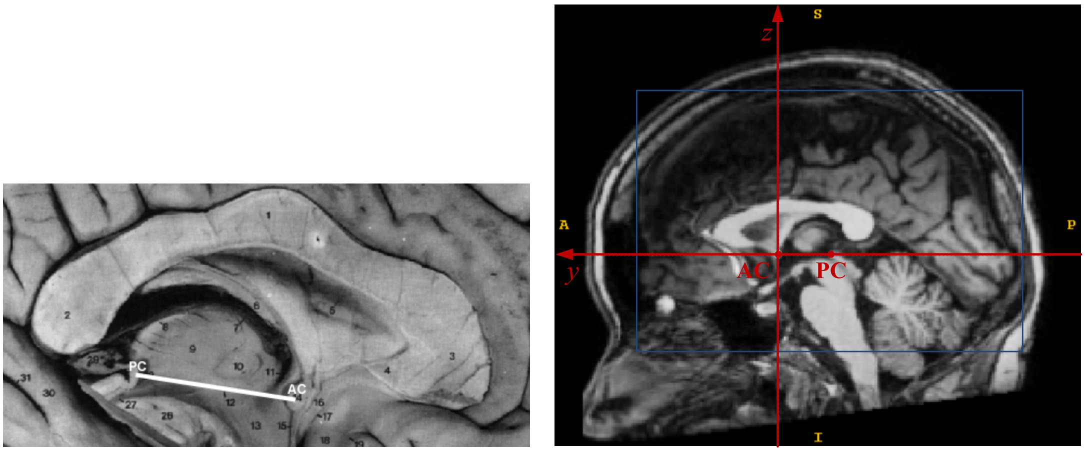
Fig. 3.7(a) and Fig. 3.7(b)) to restore the similarity transform information. Since, in fact, the two surfaces to be registered are representing the same shape (only with different rotation, scaling and translation), a simple rigid principal component alignment method is sufficient to register these two surfaces. Firstly, we move the center of mass of the SPHARM normalized surface to coincide with the center of mass of the original surface. Next, the three principal axes (eigenvectors of the points on the surface) of both the SPHARM normalized surface and the original surface are computed and aligned respectively by rotation and scaling. The results of this registration step is shown in Fig. 3.10.

After all the steps discussed above, all the procedures of mesh preparation has been finished. The final results are illustrated in Fig. 3.11. As we can see that the surface meshes have acquired the desired three properties: 1) re-meshed with required connectivity and ready to perform the wavelet transformation; 2) correspondence between each other established; 3) similarity transform information is restored.

With the mesh preparation method discussed above, in next section, a statistical surface wavelet shape model of caudate nucleus is built.

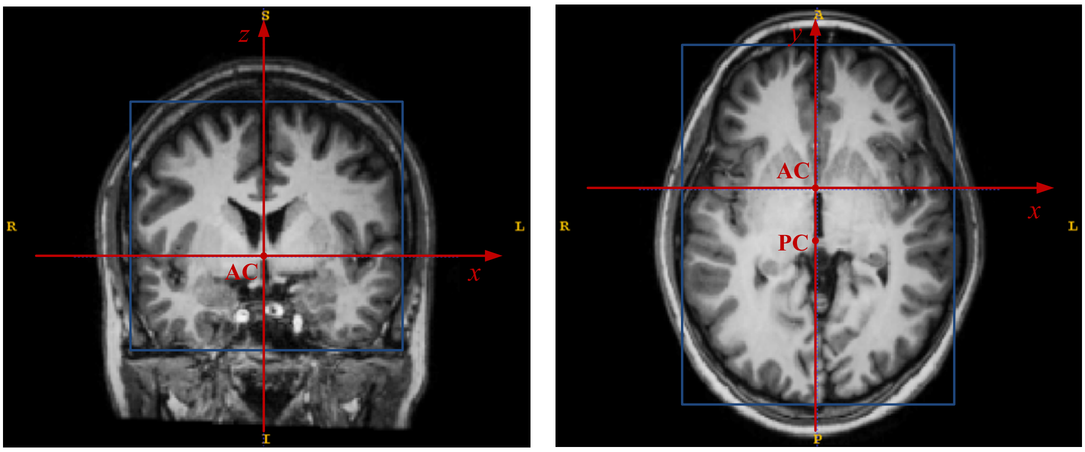
3.3 The Training of Statistical Surface Wavelet Model

In this section, we will build a statistical surface wavelet shape model of caudate nucleus based on 18 samples of caudate nucleus available from the Internet Brain Segmentation Repository (IBSR) [8]. Fig. 3.12 shows the 18 samples of caudate nucleus. Note that the shape has been normalized by the method provided in Section 3.2.



(a) Position of the anterior commissure (AC) and the posterior commissure (PC) on a post mortem mid-sagittal brain slice.

(b) sagittal view



(c) coronal view

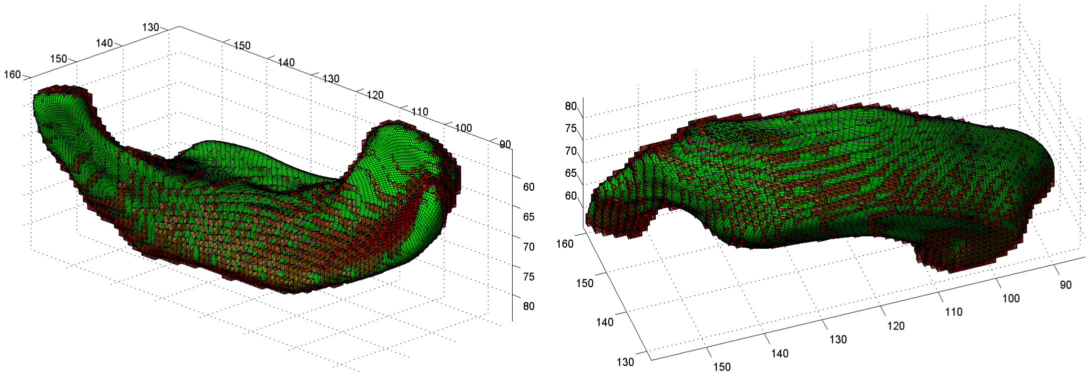
(d) axial view

Fig. 3.9: The Talairach coordinate.

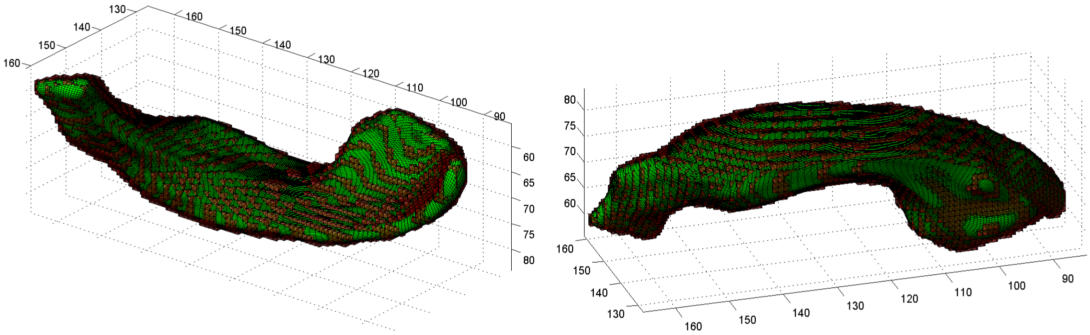
3.3.1 Decompose the Shapes in the Training Set

Having solved the correspondence issue and obtained the required subdivision mesh by re-meshing, we can now perform wavelet analysis on the surface mesh directly with the scheme introduced in (3.1). The output is a serial of wavelet coefficient vectors $\mathbf{v}_n^0$ and $\mathbf{w}_m^j$.

For our training set from IBSR, after decomposing the prepared surface meshes in the training set with this wavelet scheme, we found that $|\mathbf{w}_m^j| < 0.5$mm for $j \geq 3$. However, the minimum voxel size in IBSR is $0.8371 \times 0.8371 \times 1.5$mm,

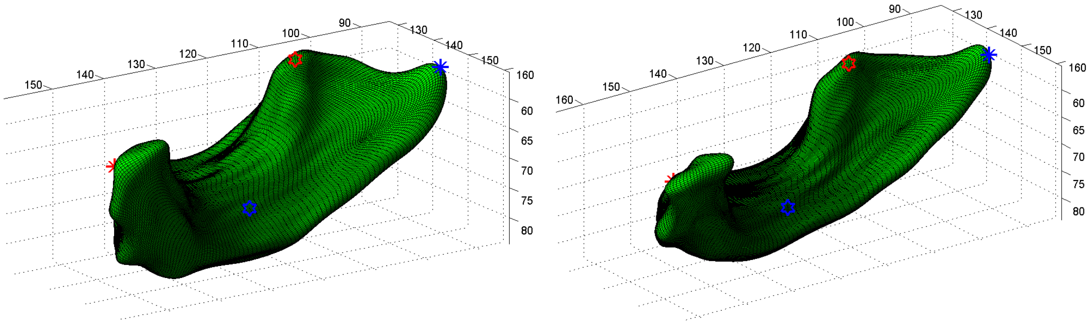


(a) cerebral lateral ventricle 1, view 1 (b) cerebral lateral ventricle 1, view 2



(c) cerebral lateral ventricle 2, view 1 (d) cerebral lateral ventricle 2, view 2

Fig. 3.10: Registration to restore the similarity transform information. (a) and (b): the registration results of cerebral ventricle 1 shown in 2 views. (c) and (d): the registration results of cerebral ventricle 2 shown in 2 views. The green mesh is the re-meshed surface with correspondences. The red mesh is the surface of the original segmented binary image data (previously shown in Fig. 3.7(a) and Fig. 3.7(b)).



(a) prepared surface mesh of ventricle 1 (b) prepared surface mesh of ventricle 2

Fig. 3.11: The re-meshed surfaces with correspondence and similarity transform information. The corresponding points are labeled with the same markers.

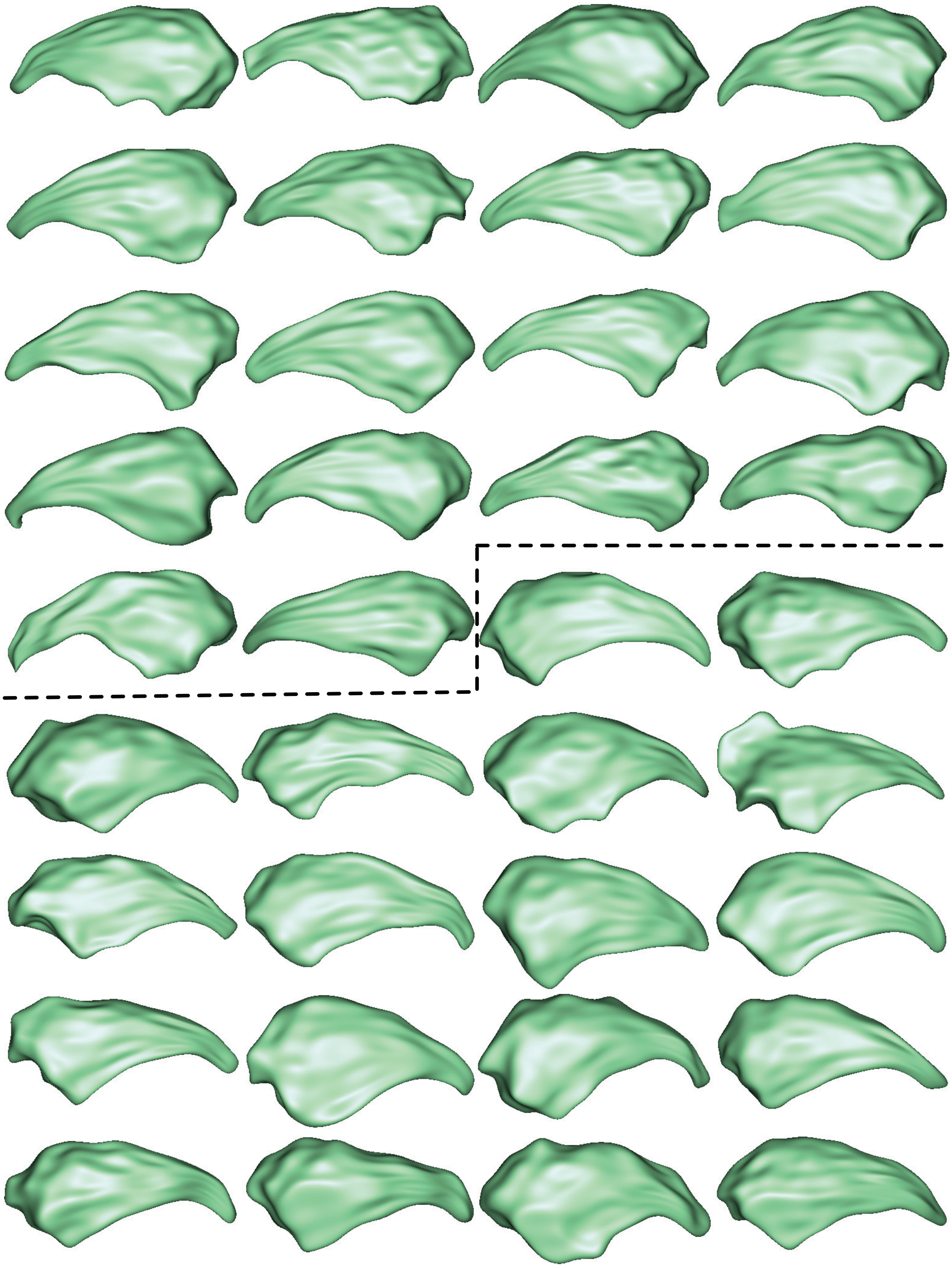


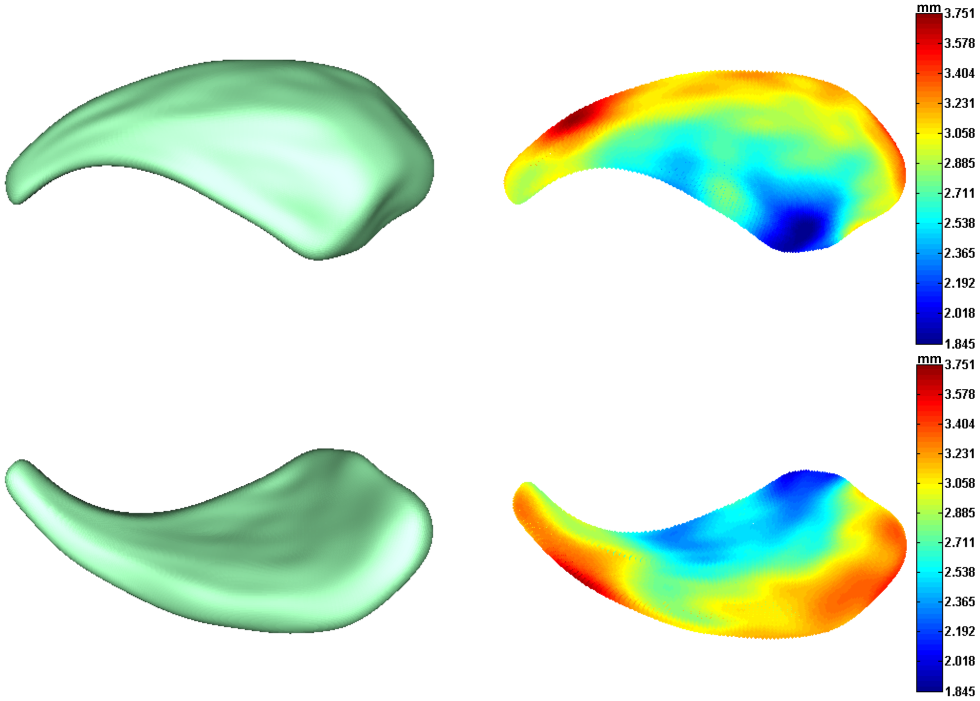
Fig. 3.12: The 18 samples of the caudate nucleus (normalized) from the Internet Brain Segmentation Repository (IBSR). Above the dashed line: left caudate nucleus; Below the dashed line: right caudate nucleus

and we can thus omit $\mathbf{w}_m^j$ for $j \geq 3$. Hence, we only use the first 4 scale levels, for a total of $|\mathbf{V}(0)| + |\mathbf{W}(0)| + |\mathbf{W}(1)| + |\mathbf{W}(2)| = 8 + 18 + 72 + 288 = 386$ coefficient vectors to describe each shape. Suppose we have N shapes in the training set. With the wavelet decomposition method described above, the i^{th} shape of the training set can be represented as: $\mathbf{C}_i = [\mathbf{c}_i^1, \cdots, \mathbf{c}_i^M]^\top$. Here, $\mathbf{c}_i^j$, $i = 1, ..., N$, and $j = 1, ..., M$ stands for the j^{th} coefficient vector of the i^{th} shape. In our case, $N = 18$ and $M = 386$.

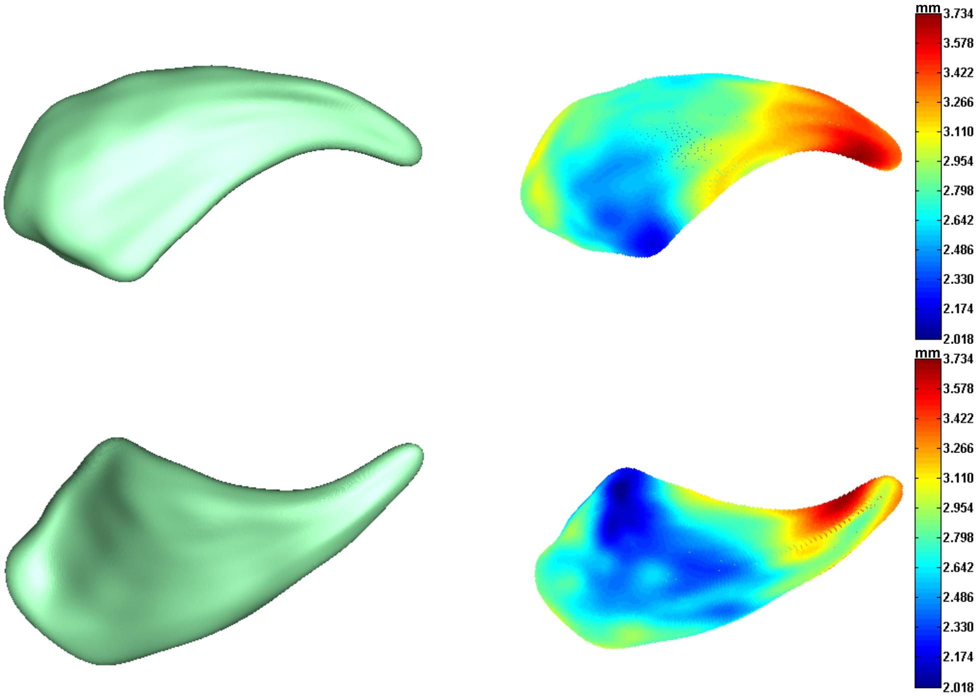
3.3.2 Computing the Statistical Surface Wavelet Model

After decomposing the training set and using $\mathbf{C}_i$, $i = 1, ..., N$ to represent the i^{th} shape, we can first compute the mean shape as: $\overline{\mathbf{C}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{C}_i$. Fig. 3.13 shows the mean shape $\overline{\mathbf{c}}$ of our training set. The mean shape will later be used as the initial guess when performing segmentation of an unknown input image. Also, the spatial distribution of the shape deviations from the mean shape can be computed. This is shown in Fig. 3.13 with different colors representing the different extent of deviations. From this figure, we can see that the tail of the caudate nucleus is the most variable part, while the middle body part is relatively stable.

Beside the mean shape, we can perform statistical shape analysis at different scales since the shape model is multiscale. By performing a principal component analysis on each specified scale level, we can understand the most significant variation of the shape in that scale level as shown in Fig. 3.14 and Fig. 3.15. Note that the variation scale changes from coarse to detail from Fig. 3.14(a) to Fig. 3.14(d) and from Fig. 3.15(a) to Fig. 3.15(d). The ASM model [2], because of its discrete shape representation using PDM (points distribution model), can't provide this kind of multiscale shape analysis capability. Models based on the Fourier [3, 10] or spherical harmonics [4, 5] basis can produce similar results as shown in Fig. 3.14 and Fig. 3.15. However, due to their global supported basis function, they don't have the ability to focus the shape analysis only on part of the surfaces at every



(a) Left caudate nucleus. upper row: view 1, lower row: view2, left: the mean shape, right: color coded shape variation distribution



(b) Right caudate nucleus. upper row: view 1, lower row: view2, left: the mean shape, right: color coded shape variation distribution

Fig. 3.13: Mean shape and shape variation distribution: (a) left caudate (b) right caudate.

scale. In contrast, because of the spatial-localization property of SSWM, we now can go further to perform shape analysis not only at a specified scale level, but also at a specified location on the surface. To do so, using the SSWM, the method is quite straightforward. Instead of doing a principal component analysis on all the wavelet coefficients corresponding to the specified scale level, we only perform the analysis on the wavelet coefficients of the specified scale level corresponding to some region of interest on the surface.

Because of the multiscale and spatial-localization property of wavelet basis, every wavelet coefficient vector only defines the shape at a specified scale and spatial location. Therefore, statistical analysis can be applied on every coefficient vector separately to obtain the multiscale and spatially localized statistical information. The 3 elements in the coefficient vector $\mathbf{c}_i^j$ represent the shape variations in the directions of the coordinate axes x, y, and z. However, the directions x, y and z do not necessarily indicate the most significant modes of the shape variation. Therefore, we use principal component analysis (as discussed in the next paragraph) to find the principal shape variation mode defined by each coefficient vector.

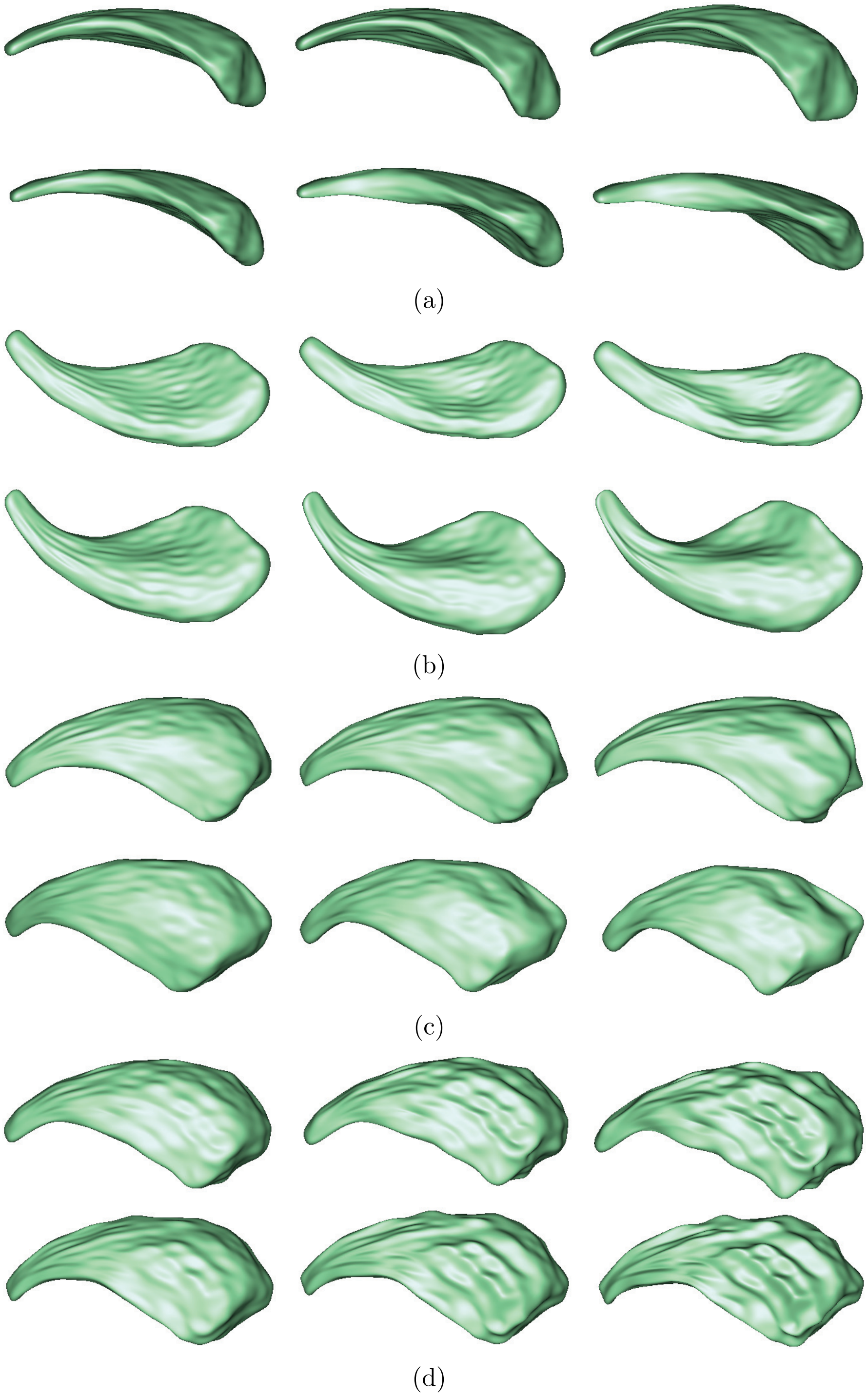
Let $\mathbf{C}^j = [\mathbf{c}_1^j, \cdots, \mathbf{c}_N^j]^\top$ be the corresponding j^{th} wavelet coefficients from the N samples in the training set. $\overline{\mathbf{c}^j} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{c}_i^j$ is the mean of the j^{th} coefficient vector. Then the covariance matrix of the j^{th} coefficient vector is:

$$\Sigma^j = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{c}_i^j - \overline{\mathbf{c}^j}) \cdot (\mathbf{c}_i^j - \overline{\mathbf{c}^j})^\top \tag{3.7}$$

Eigenanalysis of this covariance matrix produces the eigenvalues and eigenvectors. These eigenvectors define a new orthogonal basis. Using this basis, the j^{th} coefficient vector of a new shape can be expressed as:

$$\mathbf{c}^j = \overline{\mathbf{c}^j} + \mathbf{U}^j\mathbf{b}^j \tag{3.8}$$

where $\mathbf{U}^j$ is a 3×3 matrix in which the columns are the eigenvectors which



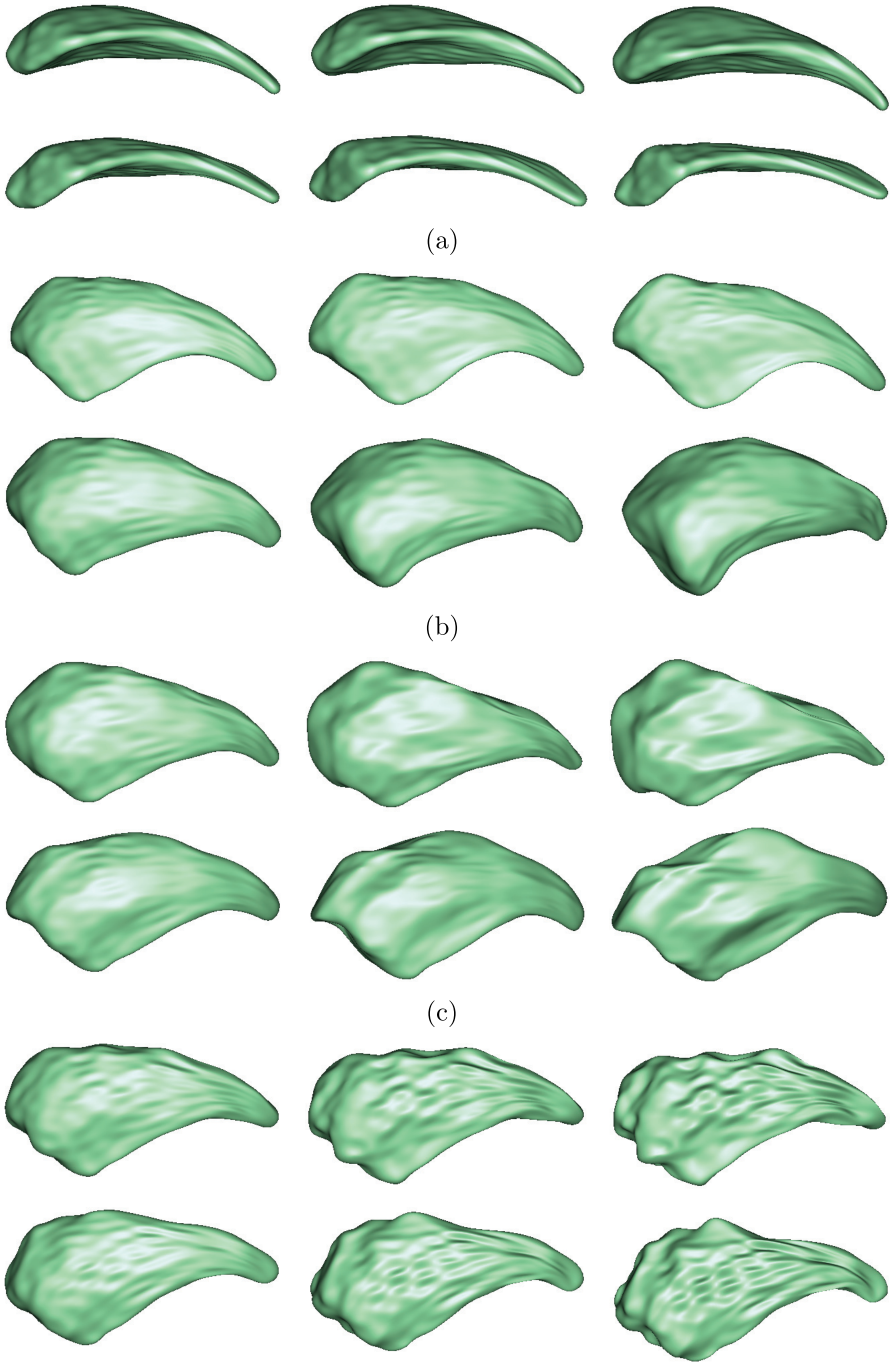
(a)

(b)

(c)

(d)

Fig. 3.14: The most significant variation mode of left caudate in different scale levels: (a) level 0; (b) level 1; (c) level 2; (d) level 3. Upper row from left to right: $mean{+}\sigma$, $mean + 2\sigma$, $mean + 3\sigma$. Lower row from left to right: $mean - \sigma$, $mean - 2\sigma$, $mean - 3\sigma$. ($mean$ is the mean shape. σ is the standard shape deviation away from the mean shape at the specified scale level.)



(a)

(b)

(c)

(d)

Fig. 3.15: The most significant variation mode of right caudate in different scale levels: (a) level 0; (b) level 1; (c) level 2; (d) level 3. Upper row from left to right: $mean + \sigma$, $mean + 2\sigma$, $mean + 3\sigma$. Lower row from left to right: $mean - \sigma$, $mean - 2\sigma$, $mean - 3\sigma$. ($mean$ is the mean shape. σ is the standard shape deviation away from the mean shape at the specified scale level.)

denote the eigenmodes of the j^{th} coefficient vector's variation (that is the eigenmodes of shape variation at the location and scale defined by $\mathbf{c}^j$); vector $\mathbf{b}^j = [b^j(1), b^j(2), b^j(3)]^\top$ is the coordinate of $\mathbf{c}^j$ in the new orthogonal basis. It describes the deviation of $\mathbf{c}^j$ from the mean value $\overline{\mathbf{c}^j}$. Without loss of generality, $b^j(1), b^j(2), b^j(3)$ correspond to eigenvalues $\lambda_1^j, \lambda_2^j, \lambda_3^j$ respectively in non-increasing significance of the variation. We define:

$$\mathbf{B} = [\mathbf{b}^1, ..., \mathbf{b}^M] \tag{3.9}$$

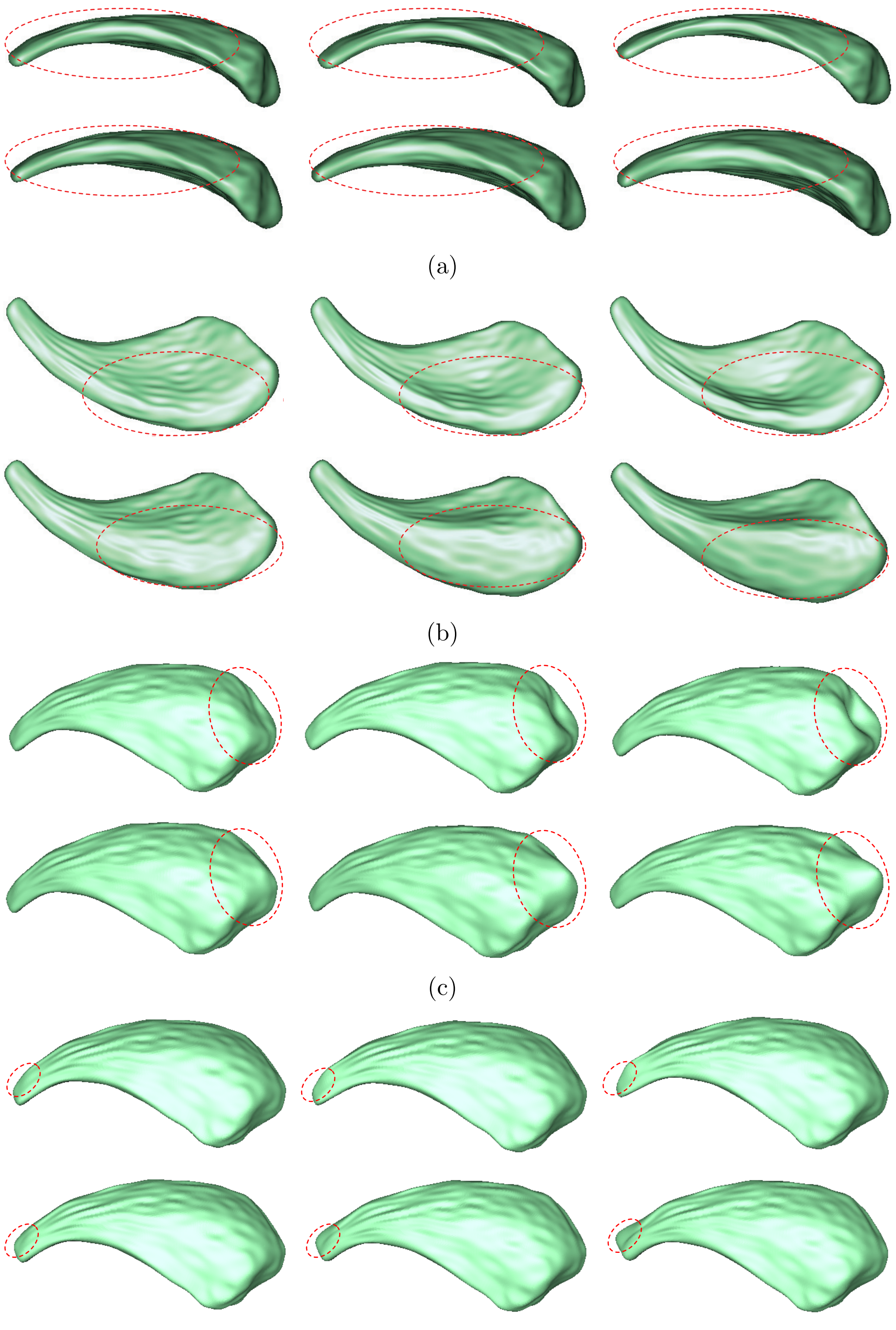
Then, matrix $\mathbf{B}$ contains parameters which describe the new shape and are, in fact, the parameters to be optimized later in the segmentation process.

Because of the randomness of the shape variation, $\mathbf{B}$ can be interpreted as random variables. However, there is no particular distribution to describe these parameters and only mean and standard deviation are known about the distribution. Because Gaussian is the natural form of a probability density and, with a given variation, the Gaussian is the one with the maximum entropy [10], a multivariate Gaussian can be assumed. A further independence assumption of this multivariate Gaussian can also been adopted, which is a common and reasonable relaxation in the parametric models (see, for example, [10] [3]). Here, the independence assumption comes from the orthogonality of wavelet basis functions and the orthogonality of the eigenvectors in the principal component analysis. With these two "orthogonality", the definition of our model does not imply any dependency between parameters $b^j(k)$. The correlations between $b^j(k)$ are, actually, additional information, and the independence assumption only ignores these additional information but not assuming more information. Because $\mathbf{b}^j$ actually indicates the deviation of the coefficient from the mean value, the mean of the assumed Gaussian distribution is 0. The standard deviation can be estimated and denoted by: $\sigma^j(k)(k = 1, 2, 3 \text{ and } j = 1, ..., M)$. Thus, the assumed independent multivariate Gaussian is defined as:

$$\Pr(\mathbf{B}) = \prod_{j=1}^{M} \prod_{k=1}^{3} \frac{1}{\sigma^j(k)\sqrt{2\pi}} e^{-\frac{(b^j(k))^2}{2\sigma^j(k)}} \tag{3.10}$$

Till here, $\overline{\mathbf{c}^j}$, $\mathbf{U}^j$, $\sigma^j(k)$, $j = 1, ..., M$ and $k = 1, 2, 3$ together constitute a statistical shape model which bias the object shape to a particular range of variation.

In order to illustrate the spatial localization property of this wavelet statistical shape model, in every scale level, we can choose one coefficient vector and show the most significant variation (corresponding to $b^j(1)$) of this chosen coefficient by changing $\mathbf{b}^j$ in (3.8). The results are shown in Fig. 3.16 and Fig. 3.17. It is very clear that when one coefficient is changed, the shape only varies locally at the specified scale, while the other parts of the shape remain unchanged. Such a multi-scale and spatial localized shape representation can be very useful in the segmentation process, because the model parameters can be optimized one by one instead of altogether at one time. This divides the complex multi-dimensional optimization problem into many small and simple one-dimension optimization problems. Therefore, this feature will lead to a more efficient model-guided segmentation method which will be discussed in next chapter.



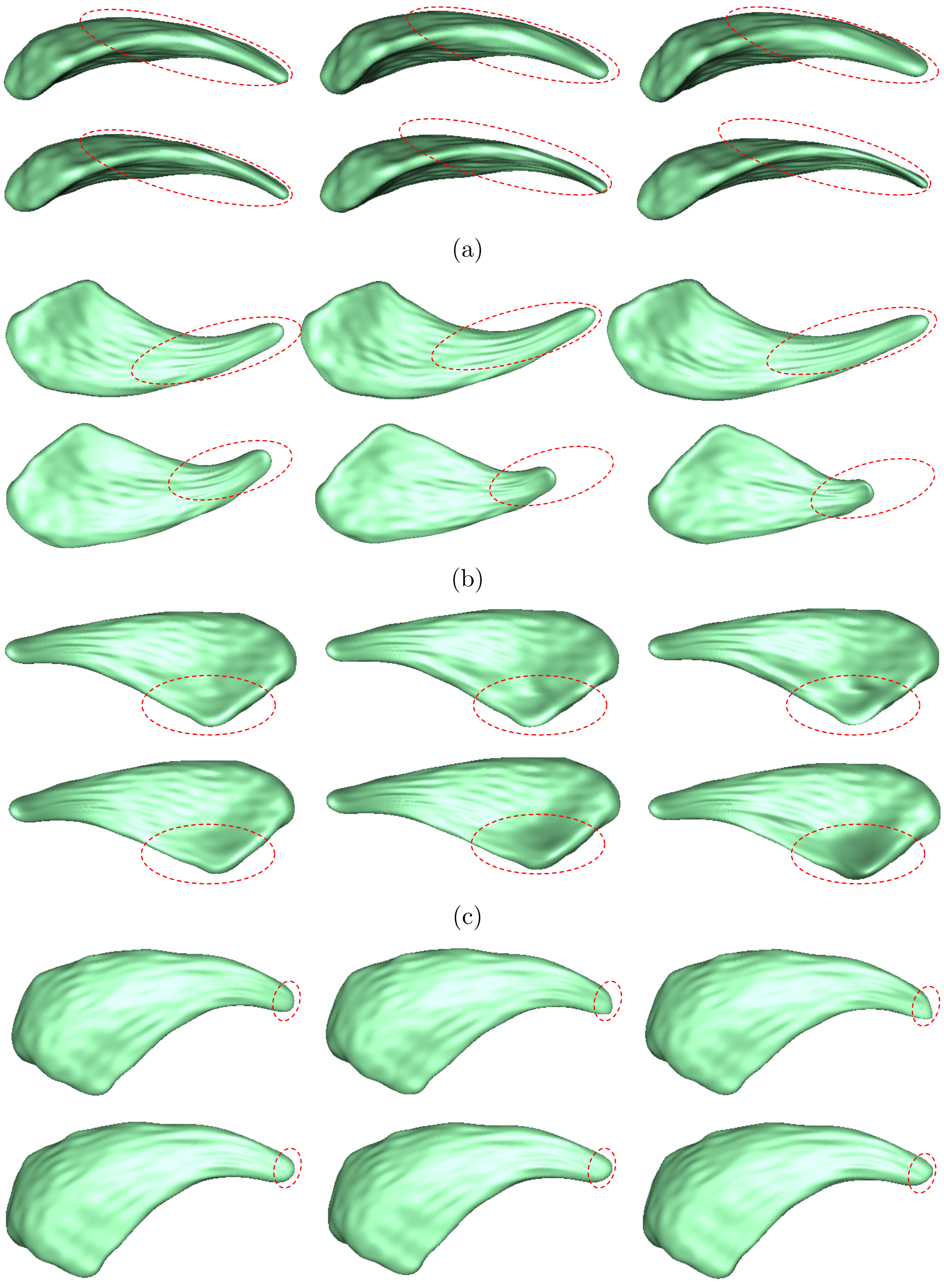
(a)

(b)

(c)

(d)

Fig. 3.16: The most significant variation mode of the left caudate nucleus at one chosen spatial location in different scale levels. (a) Level 0. (b) Level 1. (c) Level 2. (d) Level 3. The red dashed line denotes the region where the shape is defined by the chosen coefficient. Upper row from left to right: $\mathbf{b}^j = [\sigma^j(1), 0, 0]^\top$, $[2\sigma^j(1), 0, 0]^\top$, $[3\sigma^j(1), 0, 0]^\top$, respectively. Lower row from left to right: $\mathbf{b}^j = [-\sigma^j(1), 0, 0]^\top$, $[-2\sigma^j(1), 0, 0]^\top$, $[-3\sigma^j(1), 0, 0]^\top$, respectively.



(a)

(b)

(c)

(d)

Fig. 3.17: The most significant variation mode of the right caudate nucleus at one chosen spatial location in different scale levels. (a) Level 0. (b) Level 1. (c) Level 2. (d) Level 3. The red dashed line denotes the region where the shape is defined by the chosen coefficient. Upper row from left to right: $\mathbf{b}^j = [\sigma^j(1), 0, 0]^\top$, $[2\sigma^j(1), 0, 0]^\top$, $[3\sigma^j(1), 0, 0]^\top$, respectively. Lower row from left to right: $\mathbf{b}^j = [-\sigma^j(1), 0, 0]^\top$, $[-2\sigma^j(1), 0, 0]^\top$, $[-3\sigma^j(1), 0, 0]^\top$, respectively.

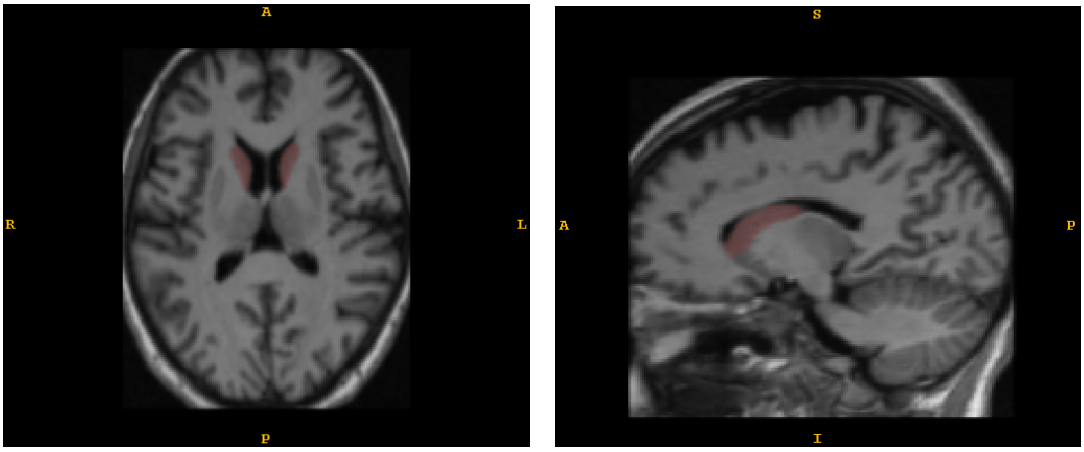
Chapter 4

SSWM-Guided Segmentation

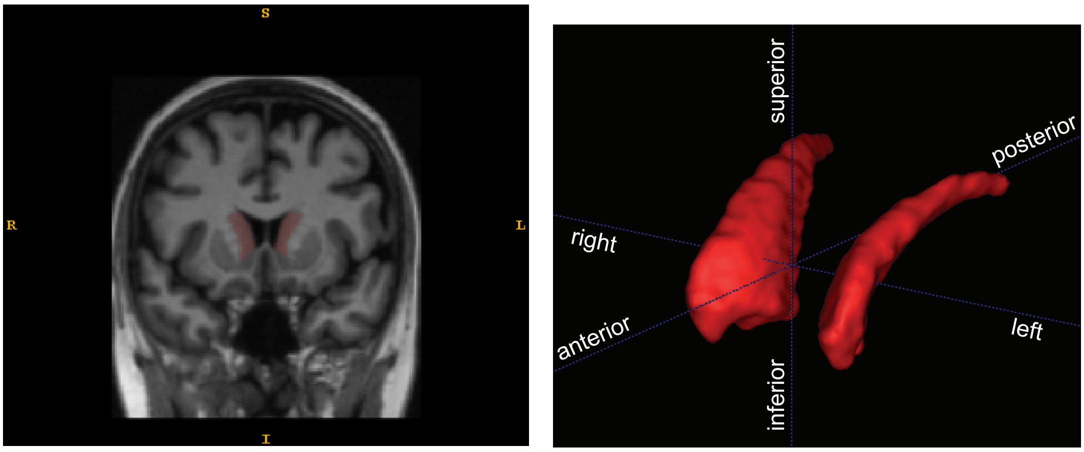
Segmentation of anatomical structures from magnetic resonance imaging (MRI) or computed tomography (CT) data sets is the first and basic step in many medical image applications, such as diagnosis, therapy evaluation, surgical planning and navigation. Because manual segmentation is time-consuming and lacks reproducibility, the development of automated or semi-automated techniques is highly desirable. In general, the development of computer-assisted segmentation methods is challenging due to many difficulties. For example, the intensity based methods are likely to fail if the distribution of intensity values of the structure of interest overlaps with those of the surrounding structures. Moreover, in medical images, the object boundaries are often smeared due to low image contrast or even missing when blended with other surrounding structures with similar intensity values. Therefore, boundary-based non model-guided methods such as "snake" [22] may "leak" and result in a poor segmentation. We use the caudate nucleus as an example to illustrate this difficulty in segmentation. The caudate nucleus is a nucleus located within the basal ganglia of the brain (see Fig. 4.1), originally thought to primarily be involved with control of voluntary movement and is now known to be an important part of the brain's learning and memory system [73]. However, some part of the boundary of the caudate is ofter smeared by the surrounding gray matter structures, the putamen and the accumbens-area in MR images (as shown in Fig. 4.2). Therefore, the model-free techniques often cause serious leakage at these areas.

To overcome these difficulties, a model with prior knowledge of the target object can be very helpful. A number of statistical shape models as indicated in Chapter 2 have been proposed for this purpose. Our proposed SSWM can also be used in the model-guided segmentation. Moreover, because of its superior "double localization" property, our model will result in a multiscale and spatial localized optimization algorithm. This is a valuable advantage over other models, because, as we know, detailed contours and especially surfaces often require a large number of coefficients to capture and these coefficients need an optimization process to determine. Hence, the dimensionality of the optimization problem is usually of the order of several tens to a few hundreds, with the latter to be more realistic. Not every optimization method is able to handle that many free parameters. Possible choices reach from deterministic techniques like conjugate gradient optimization to stochastic methods such as genetic algorithms. For the gradient-based algorithms, they are often sensitive to the initial location because of the many local minimals. The stochastic methods are usually quite time consuming and can only converge and find the right minimal at certain probability. However, with our new SSWM model, this multivariant optimization problem in model-guided segmentation can be divided into many univariate problems and solved in a much more efficient manner.

Having built the statistical surface wavelet shape model of the caudate nucleus in Chapter 3, next in this chapter, we will perform model-guided segmentation by deforming the model to fit with the input image. This deformation process is driven by the optimization (Section 4.2) of an objective function (Section 4.1).



(a) axial (b) sagittal



(c) coronal (d) 3D visualization

Fig. 4.1: The caudate nucleus shown in axial, sagittal and coronal slices of a MR image.

4.1 The Segmentation Objective Function

Let $\mathbf{B} = [\mathbf{b}^1, ..., \mathbf{b}^M]$ be a shape template in the form of (3.9) and $G(x, y, z)$ be an input image. In order to segment the object from the image, some preprocessing of the image is needed. We calculate a Canny gradient magnitude image (as shown in Fig. 4.4(a)) after applying a Gaussian smoothing ($\sigma = 3$ voxels). The reason of performing the smoothing is to suppress the noise and to increase the capturing range. Next, this gradient magnitude image is also normalized to an intensity value range of $[0, 1]$. Let $I(x, y, z)$ denote the preprocessed image (such as the one

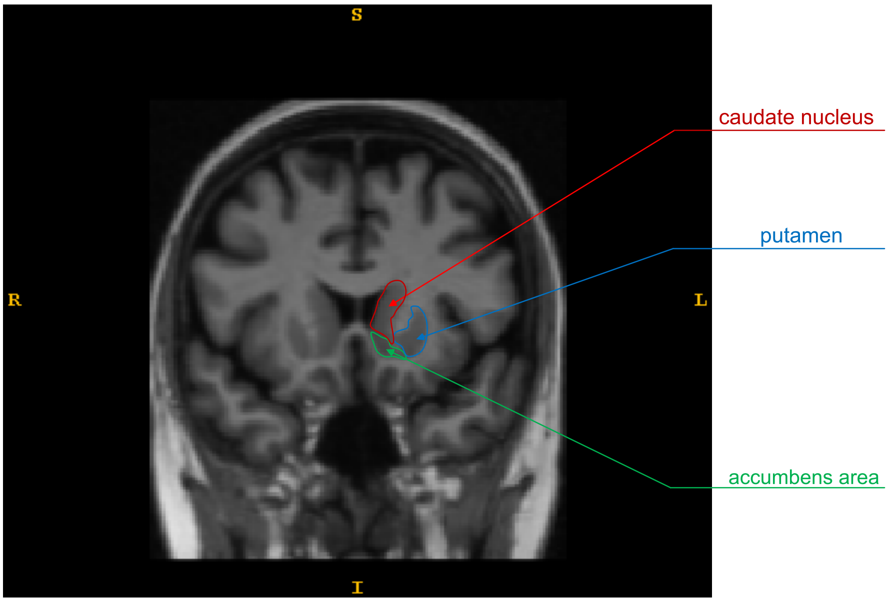


Fig. 4.2: The difficulties of segmentation of caudate nucleus

shown in Fig. 4.4(a)) in which a bigger intensity value indicates a stronger edge feature in the original image $G(x, y, z)$.

The model-guided segmentation problem can be formulated as an optimization problem:

$$\max_{\mathbf{B}} E(\mathbf{B}, I) = \max_{\mathbf{B}} [E_I(\mathbf{B}, I) + \alpha E_p(\mathbf{B})] \tag{4.1}$$

The first term $E_I(\mathbf{B}, I)$ is the image-driven term that measures the fitness between the shape template $\mathbf{B}$ and image I. It drives the deformation of the shape template to match better with the edge information in I. It is computed as:

$$E_I(\mathbf{B}, I) = \sum_{\mathbf{x} \in \mathcal{A}} I(\mathbf{x}) \Delta\mathbf{x} \tag{4.2}$$

where $\mathcal{A}$ (see Fig. 4.3) is the surface defined by $\mathbf{B}$; $\mathbf{x}$ is a point on the surface $\mathcal{A}$; $\Delta\mathbf{x}$ is the surface element (see Fig. 4.3); and $I(\mathbf{x})$ is the edge map intensity at point $\mathbf{x}$. The second term $E_p(\mathbf{B})$ is the prior information term which drives the template deformation to the most probable shape according to the prior knowledge. Since

$\mathbf{B}$ is the assumed independent multivariate Gaussian, the second term can be defined as:

$$E_p(\mathbf{B}) = \ln\Pr(\mathbf{B}) \tag{4.3}$$

Substituting $\Pr(\mathbf{B})$ in (3.10) into (4.3) gives:

$$\begin{aligned} E_p(\mathbf{B}) &= \ln\left(\prod_{j=1}^{M}\prod_{k=1}^{3}\frac{1}{\sigma^j(k)\sqrt{2\pi}}e^{-\frac{(b^j(k))^2}{2\sigma^j(k)}}\right) \\ &= \sum_{j=1}^{M}\sum_{k=1}^{3}\left[\ln\left(\frac{1}{\sigma^j(k)\sqrt{2\pi}}\right) - \frac{(b^j(k))^2}{2\sigma^j(k)}\right] \end{aligned} \tag{4.4}$$

Note that α in the objective function is a weighting factor to balance the importance between $E_I(\mathbf{B}, I)$ and $E_p(\mathbf{B})$. For input image of low contrast, a large α is preferred to bias segmentation towards prior model, while for input image of good contrast, a small α is suffice to rely more on the image information. In practice, α is to be calibrated first through experimentation on images obtained from the same source.

4.2 Optimization of the Objective Function

In general, the optimization problem in model-guided segmentation is a global optimization problem. Due to the existence of edge features of surrounding objects, the objective function has many local maximals. One needs to ensure that the global maximum is close to the initial guess of the solution in order to use local optimization algorithms, such as gradient descent, direction set and conjugate gradient [53][10][3][24]. On the other hand, with our wavelet model, this is not much of an issue. Because of the localization property of the wavelet model both in frequency and spatial location, the objective function can be optimized in a serial fashion—every time only one parameter from $\{b^j(k) | j = 1, ..., M; k = 1, 2, 3\}$

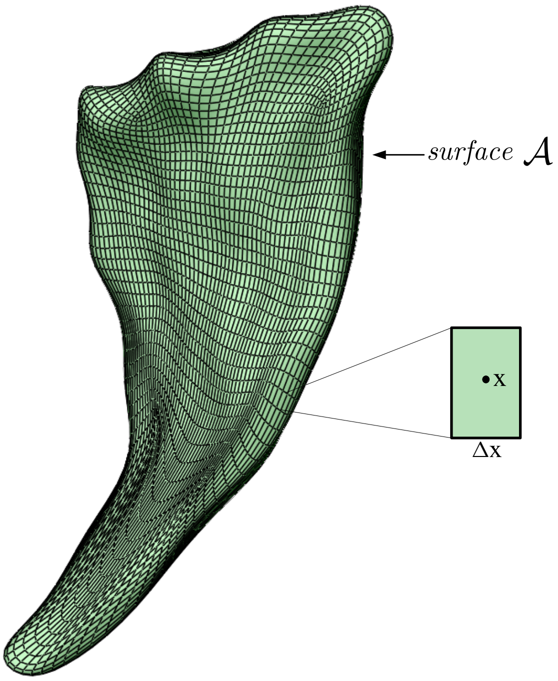


Fig. 4.3: The surface $\mathcal{A}$ and the surface element.

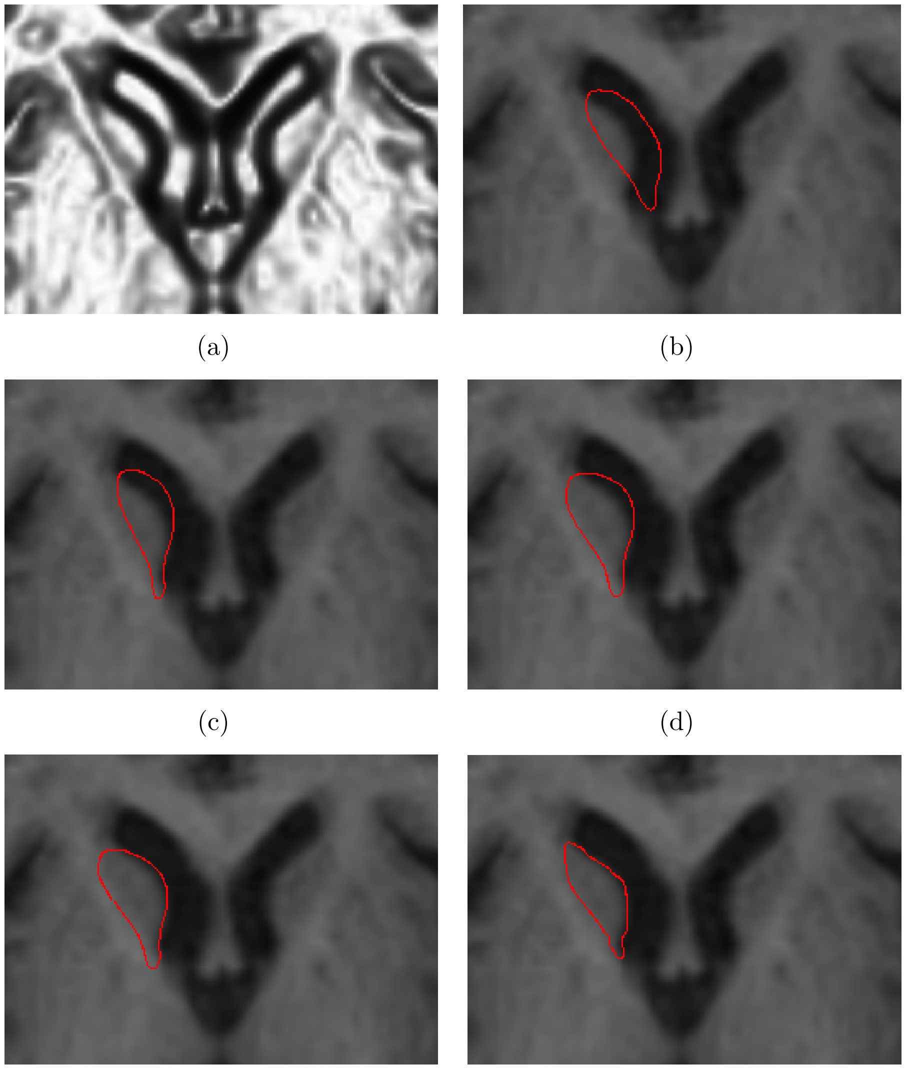
needs to be optimized. This means that there is only one 1-dimensional optimization problem to be solved each time. Therefore, we adopt the equal sampling method to solve this relatively simple 1-dimensional optimization problem. For every parameter $b^j(k)$, within the ±2 standard deviation, the objective function is evaluated at equidistant sampling points to find the maximum. The larger the number of sampling of points, the more accurate the maximum can be found. The number of sampling points is decided proportionally by the width of the searching scope defined by ±2 standard deviation. For example, if the range defined by the ±2 standard deviation is 10mm wide and the resolution of the original image is 1mm, 10 sampling points (1mm spacing between each other) will be enough.

When there are p parameters to be optimized in one scale level, and for every parameter there are l number of sampling points, the computational complexity of our optimization in this scale level is then $O(pl)$. In contrast to other shape models with only frequency localization (such as [3][4]), an equal sampling optimization method has, in general, computational complexity of $O(l^p)$. This is because in these cases, all the parameters (which are correlated to each other) in a scale level have to be optimized altogether. Here, the number of parameters p in one scale level is usually tens or hundreds, for these models, such an optimization method is thus computationally expensive and may be even unaffordable.

4.3 The Segmentation Results

We have applied our method to the segmentation of cerebral caudate nuclei. The 18 samples in IBSR [8] are used as the training set to obtain a statistical surface wavelet model with 4 scale levels (as shown in Chapter 3). To segment an input image, we first locate the Talairach landmarks manually to define the Talairach coordinate system for the input image. Next, the model to segment the input image is initialized to the mean shape (Fig. 3.13) at the start of the optimization (as shown in Fig. 4.4(b)). Note that in this example in Fig. 4.4, the initial position is quite far away from the target, but close to the “misleading” edge of lateral ventricle. Each successive parameter optimization is done in a multiscale manner starting from the coarsest level to the finest level. Within every level, the parameters in this level are optimized sequentially one by one. Because of the spatial localization of the wavelet model, optimization of one parameter only results in the deformation of a subpart of the surface, without affecting previously fitted parts. So, a sequential approach can fit the whole model to the input image, scale by scale and part by part. This is clearly visible in Fig. 4.4(b)-Fig. 4.4(e). Fig. 4.4(f) shows the final segmentation result after optimization in 4 scales in an axial 2D slice, where more detailed shape information is found. In Fig. 4.5 and Fig. 4.6, another example of segmentation process are shown in 3D in two views (superior and lateral). From the figures, we can see clearly how the model deform gradually scale by scale and part by part to fit with the target shape.

In our experiment, we also observe the usefulness of a prior model such as our statistical surface wavelet model. For the coronal slice shown in Fig. 4.10,

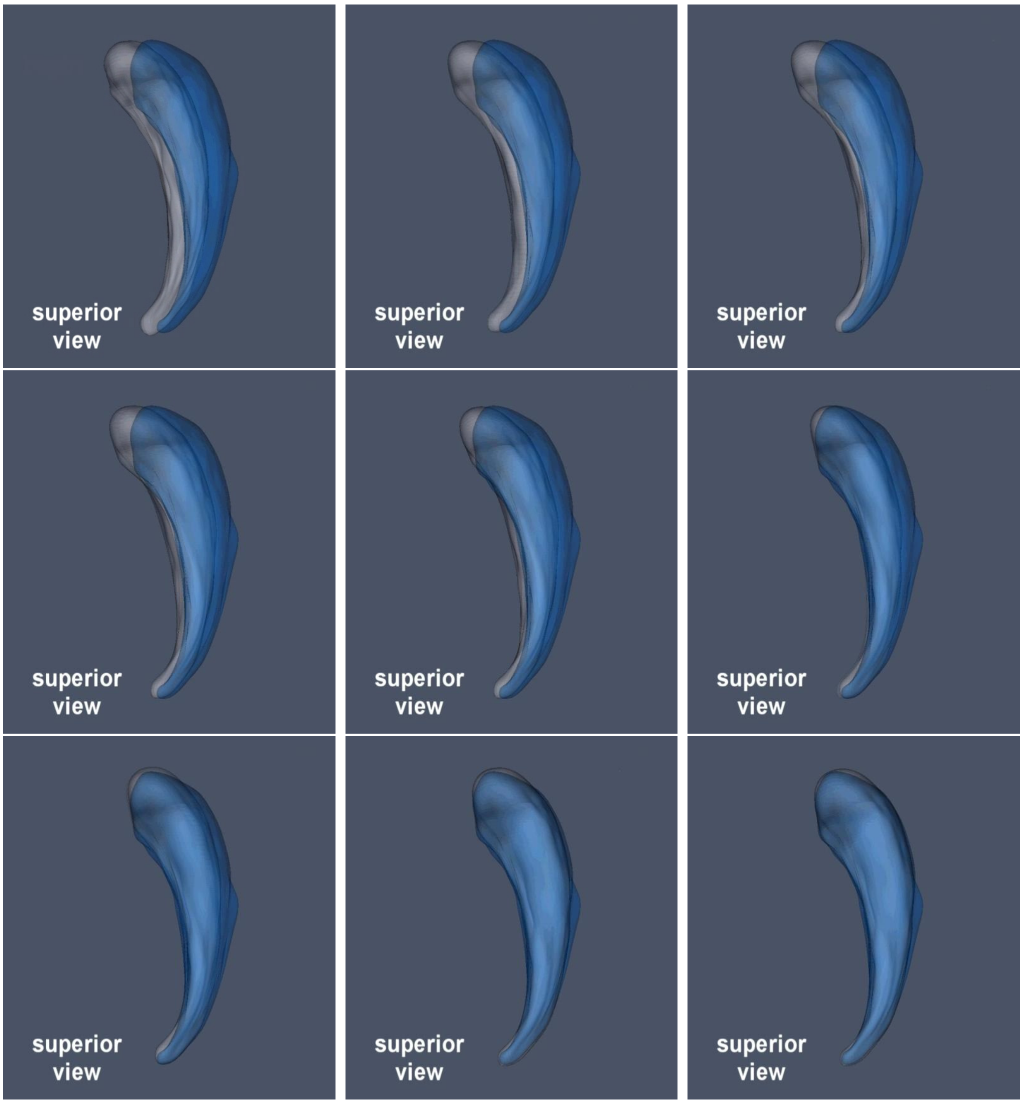


(a) (b) (c) (d) (e) (f)

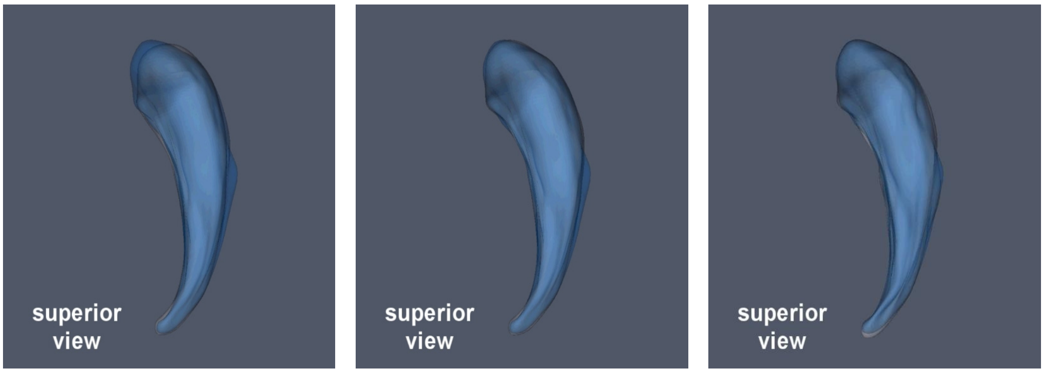
Fig. 4.4: The model deformation process shown in axial 2D intersections at the coarsest level. (a) The preprocessed image. (b) The model initialization. (c)-(e) 3 interim steps of optimization at scale level 0. (f) Final result after optimization up to scale level 3.

the caudate blends with the putamen and accumbens-area and there is thus no reasonable "edges" detected among them. It is highly improbable to obtain a good segmentation without the guidance of a statistical model.

In order to test the generality of the model, besides 36 normal cases of MR scans, 29 additional cases with schizophrenia are also used to test our method. For all these 65 MR scans, our proposed method (implemented in C++) successfully

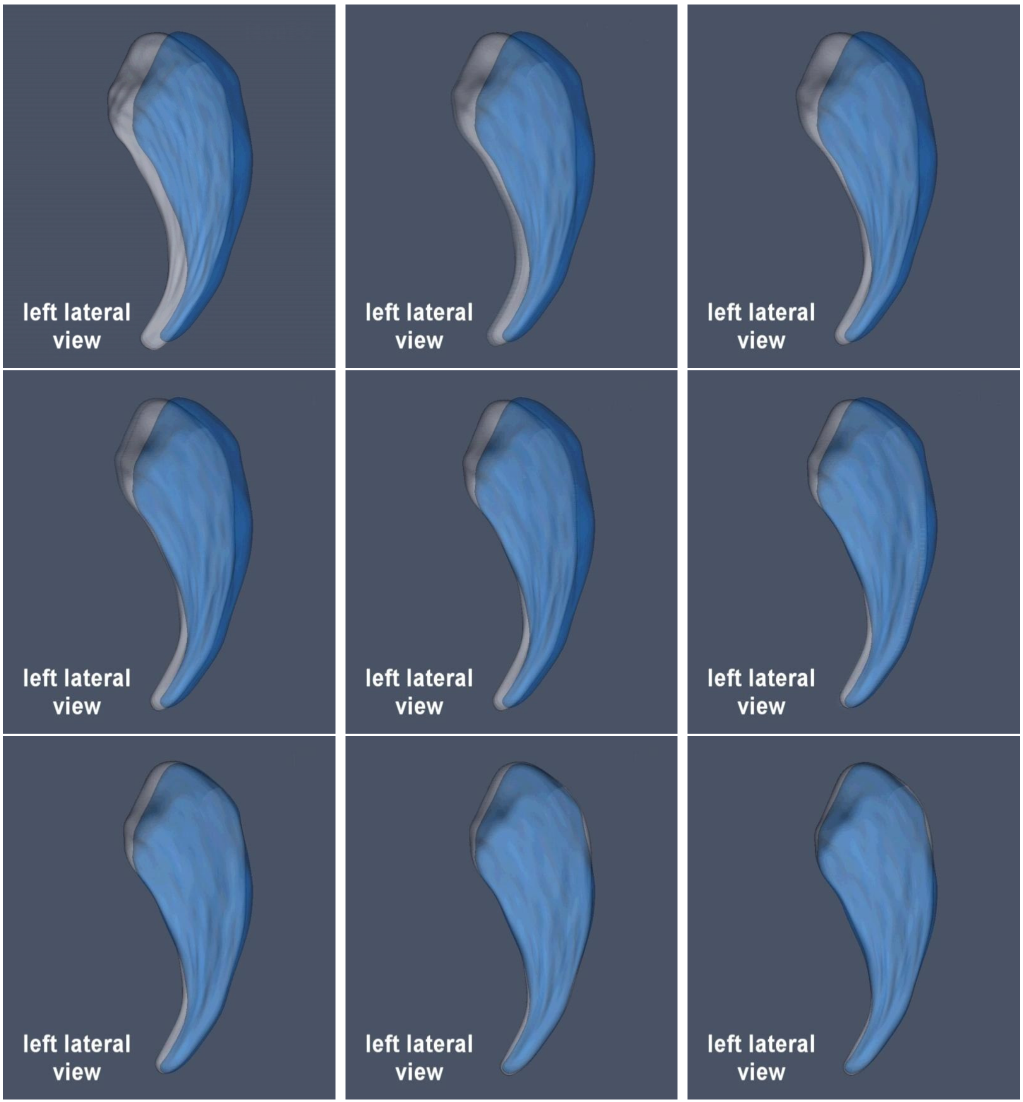


(a) deformation process in Level 0

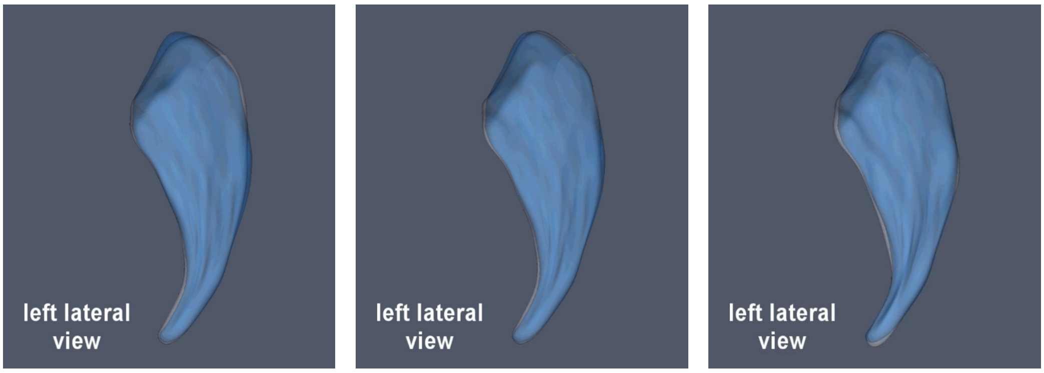


(b) result after Level 1 (c) result after Level 2 (d) result after Level 3

Fig. 4.5: The model deformation process shown in 3D at superior view. The manually segmentation is shown in light blue and the model is shown in light grey.



(a) deformation process in Level 0



(b) result after Level 1 (c) result after Level 2 (d) result after Level 3

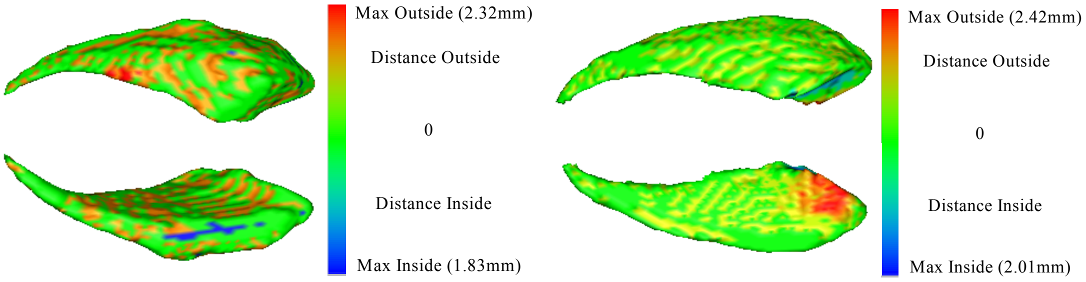
Fig. 4.6: The model deformation process shown in 3D at left lateral view. The manually segmentation is shown in light blue and the model is shown in light grey.

segmented the left and right caudate within 3 minutes on a P4 2.4GHz Windows XP system. To validate the segmentation result, we compared the results with the results obtained through manual segmentation. By using the Valmet software [74], three metrics are used to measure the segmentation errors.

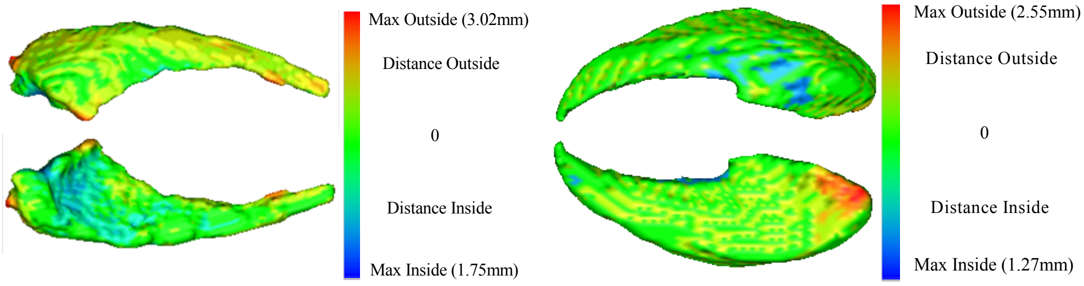
The first is the volumetric Overlap ratio. In this metric, two binary images are analyzed voxel by voxel to calculate false positives, false negative, true positive and true negative voxels. Here we accepted the measure of the intersection of two subjects and divided by the union, $(A \cap B)/(A \cup B)$. This measure gives a score of 1 for perfect agreement and 0 for complete disagreement. Because most of the error occurs at the boundary of objects, small objects are penalized and get a much lower score than large objects. Since caudate is a small structure, this is a strict metric for caudate segmentation. The second metric used is the Hausdorff distance which defines the largest difference between two objects. Given two objects A and B, the Hausdorff distance between them is defined as:

$$D_H(A, B) = \max\{\max_{p \in A}\{\min_{q \in B} d(p, q)\}, \max_{q \in B}\{\min_{p \in A} d(p, q)\}\} \tag{4.5}$$

where p and q are points on A and B, respectively. $d(p, q)$ is the Euclidean distance between p and q. From the definition, we can see that the measure is extremely sensitive to outliers and does not reflect properties integrated along the whole surface and only give the “worst case distance”. So, in order to have an overall measure of surface distance, we adopted the third metric, the average surface distance which tells how much on average the two surfaces differ. This measure integrates over both outside and inside errors on a surface and divide by the number of boundary points. This absolute average distance, as opposed to binary overlap ratio, does not depend on the object size. The validation results in the above three metrics are summarized in Fig. 4.8 (left caudate nucleus) and Fig. 4.9 (right caudate nucleus). From these figures, we can see that our proposed method is robust and achieves a high accuracy. In order to see the distribution of segmentation errors on the surface, validation results of 4 scans are shown in



(a) example 1 (b) example 2



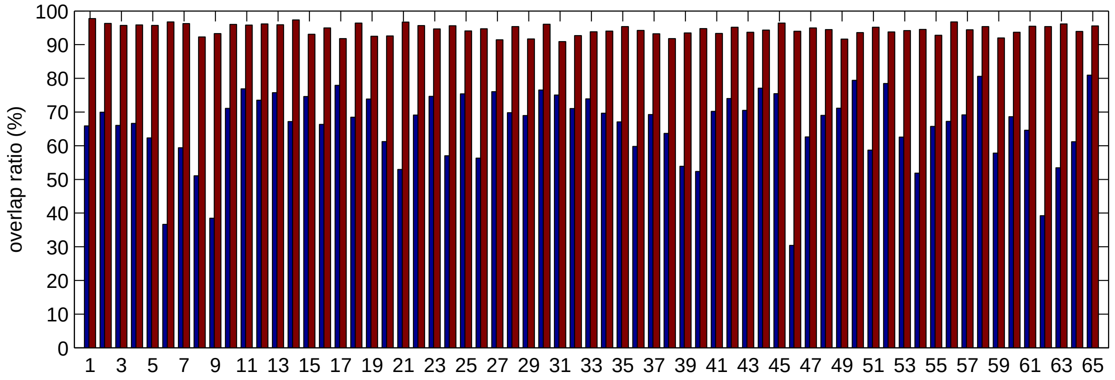
(c) example 3 (d) example 4

Fig. 4.7: Four examples of validation results shown in color-coded map. The pictures are generated by the software Valmet [74]

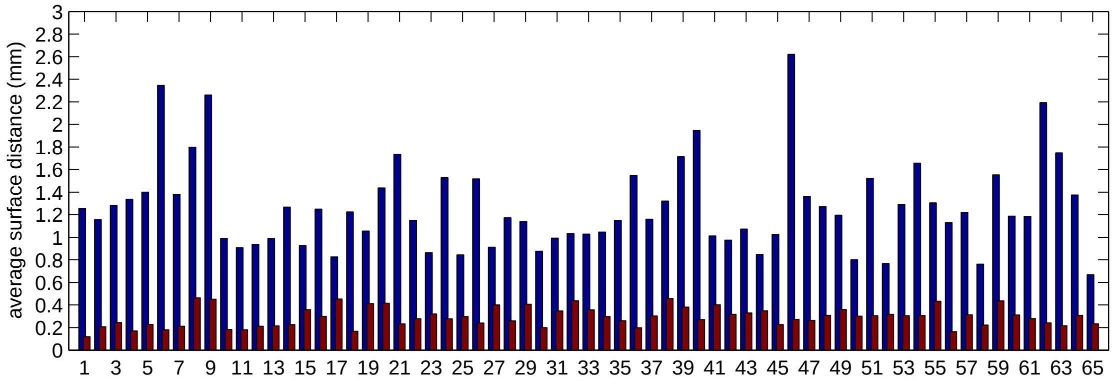
Fig. 4.7 as examples, where we can see that most of the segmentation error was generated at the area adjacent to accumbens-area.

Besides the caudate nucleus, we also test our proposed segmentation framework on putamen which is a structure in the middle of the brain (as shown in Fig. 4.2). Putamen and the caudate nucleus form the dorsal striatum. It is known that putamen plays a role in reinforcement learning [75]. For putamen segmentation, the prior model training and the segmentation process is the same as that of caudate nucleus. For both left and right Putamen, we have totally 130 cases. In most of them, we get good results. However, due to the higher variability of putamen and the more complicated surroundings of it, we encountered the following two scenarios in segmentation.

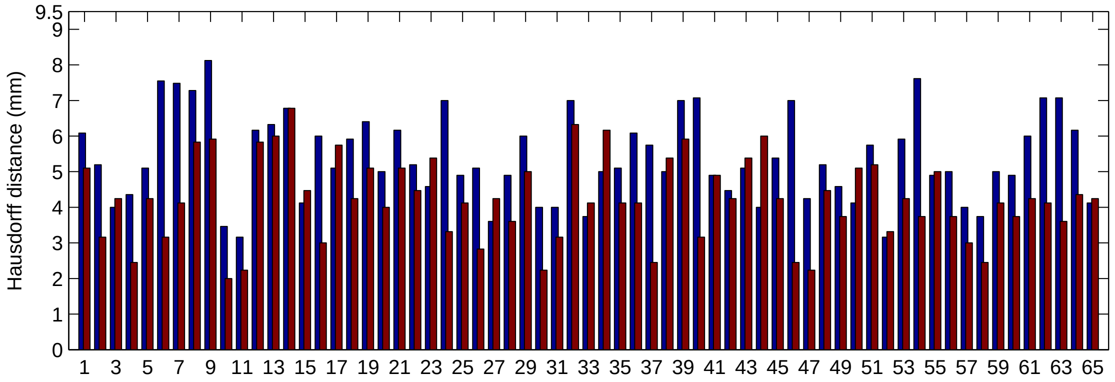
Scenario A(16 cases): The edge information is missing at some part of the boundary. Therefore, the model is attracted by the surrounding structure's stronger edge feature. Fig. 4.11(a) shows a typical edge-map in this scenario. We



(a) overlap ratio(%). Before deformation: 65.8±10.8/After deformation: 94.5±1.6



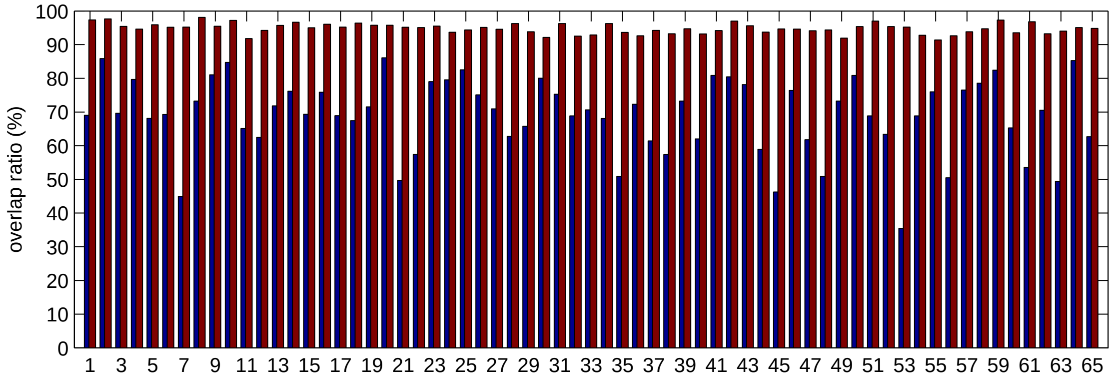
(b) average surface distance(mm). Before deformation: 1.27±0.40/After deformation: 0.29±0.09



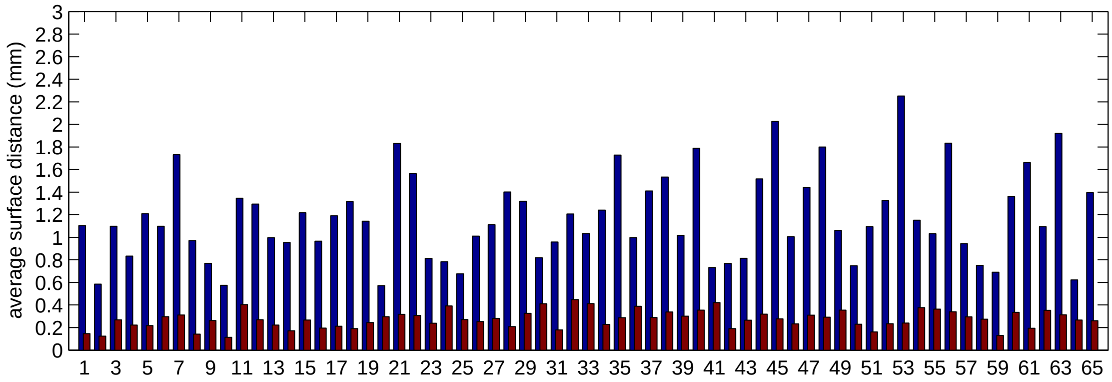
(c) Hausdorff distance(mm). Before deformation: 5.39±1.23/After deformation: 4.23±1.16

Fig. 4.8: Segmentation results of 65 left caudate. Bars in blue illustrate the measure at initialization and in red after deformation

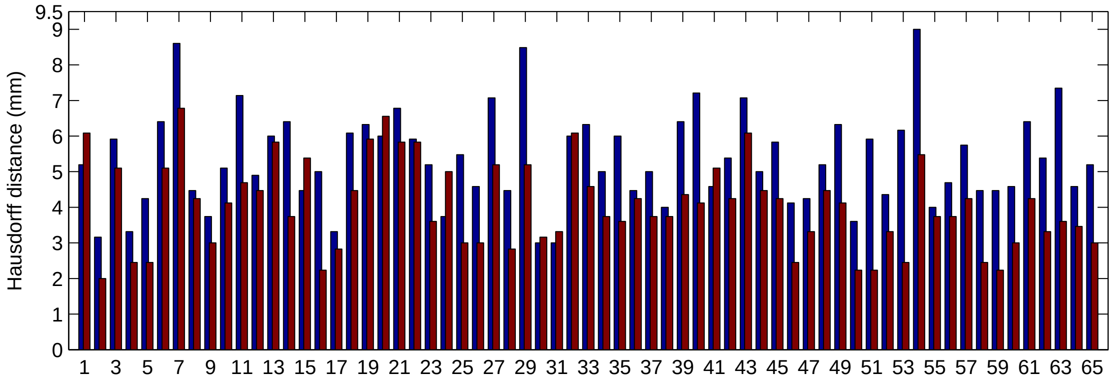
can see that part of the putamen's edge is totally invisible (the part is indicated as a red curve in Fig. 4.11(b)). However, the surrounding structure has a very strong edge feature (indicated as a green curve in Fig. 4.11(b)) and has a similar



(a) overlap ratio(%). Before deformation: 68.9±11.2/After deformation: 94.8±1.5



(b) average surface distance(mm). Before deformation: 1.17±0.39/After deformation: 0.27±0.08



(c) Hausdorff distance(mm). Before deformation: 5.35±1.33/After deformation: 4.04±1.23

Fig. 4.9: Segmentation results of 65 right caudate. Bars in blue illustrate the measure at initialization and in red after deformation

shape. Therefore, the model is attracted to this false edge and results in a bad segmentation as Fig. 4.11(c) shows. In this situation, prior knowledge can't help, because the deformation to the false edge does not exceed the deformation range (2

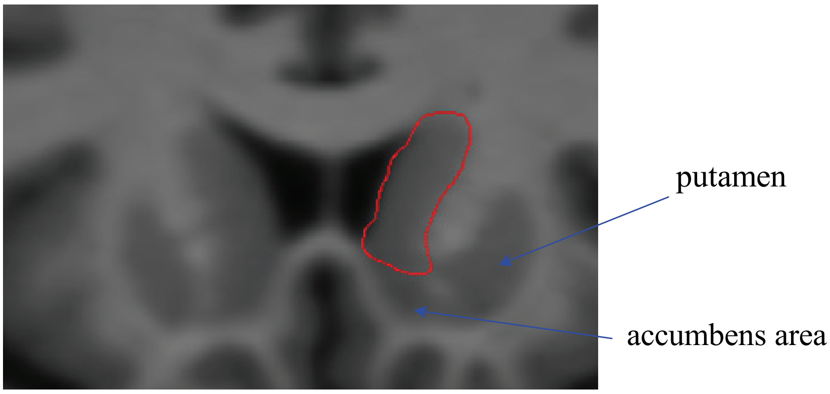


Fig. 4.10: The separation between caudate, putamen and accumbens-area using the prior knowledge.

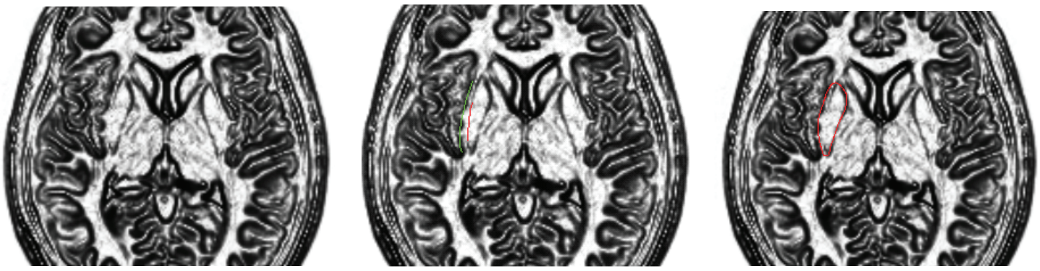
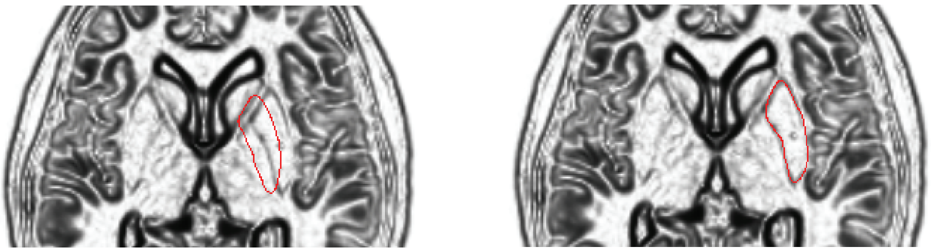

(a) an axil slice (b) the missing edge (c) the wrong result

Fig. 4.11: The scenario A in putamen segmentation, in which the edge information is missing at some part of the boundary and the model is attracted by the surrounding structure's stronger edge feature.

standard deviation) defined by the prior model. In some senses, the algorithm does nothing wrong, since the real edge feature is missing, it just finds the most probable substitute according to the prior knowledge. To summarize, the happening of this scenario has three necessities: (a) the edge of putamen is missing; (b) the false edge has the similar shape as that of the missing edge; (c) the false edge is within the range of 2 standard deviation.

Scenario B(17 cases): Fig. 4.12(a) shows an example that using the prior model trained based on the 18 IBSR [8] samples failed to segment. The reason of this failure is that the training set does not contain this pattern of shape variation. By selecting 3 cases from the 17 failed cases and adding them to the 18 IBSR



(a) the segmentation result with the old prior model (learnt from the 18 samples of IBSR)

(b) the segmentation result with the new prior model (learnt from the 18 samples of IBSR and 3 failed cases)

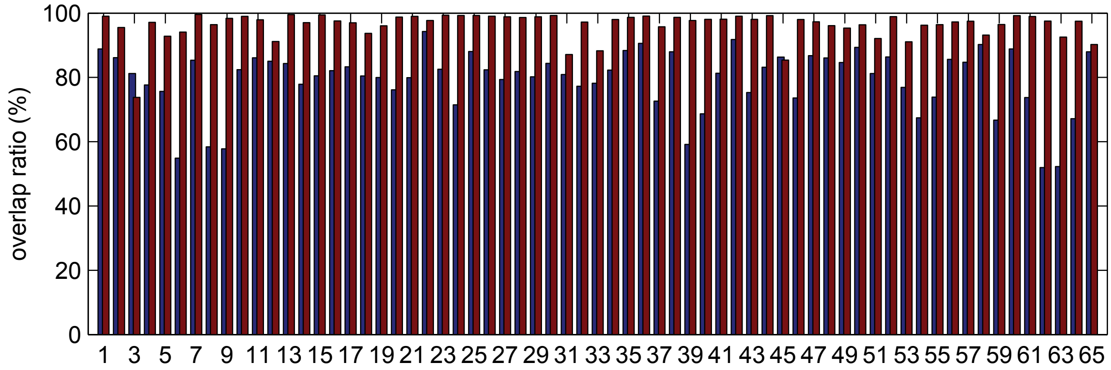
Fig. 4.12: The scenario B in putamen segmentation, which contains shape variation pattern not included in the 18 samples of IBSR.

samples to form a bigger training set, we compute a new prior model. With the new model, we can segment not only the previously succussed cases but also the 17 previously failed cases (as shown in Fig. 4.12(b)). This shows that all these 17 failed cases belong to the same shape variation pattern which is not contained in the 18 samples of IBSR. Scenario B is actually within our expectation, because it is not possible that 18 IBSR samples can contain all the possible variation patterns and, therefore, we always need to update the prior model with the newly found shape variation patterns.

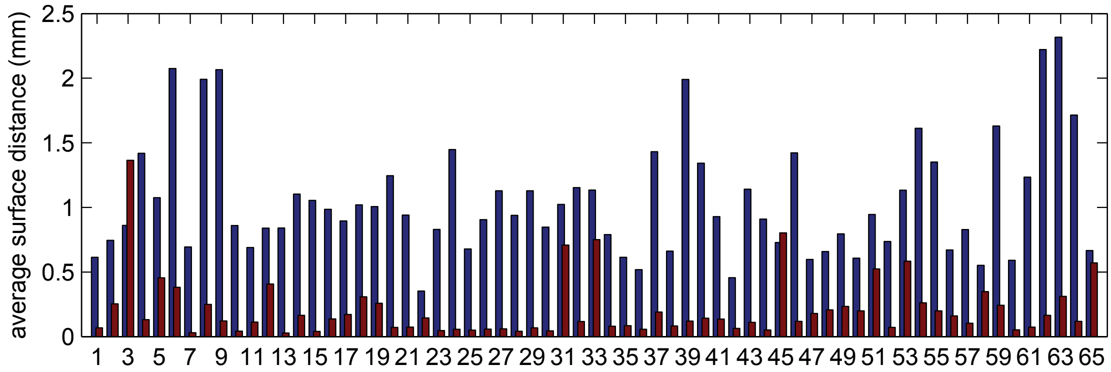
The segmentation results of left and right putamen using this updated prior model are summarized in Fig. 4.13 and Fig. 4.14, respectively. As we can see, except some cases belong to Scenario A, the algorithm generates very good segmentation results.

4.4 The SSWM Segmentation Software

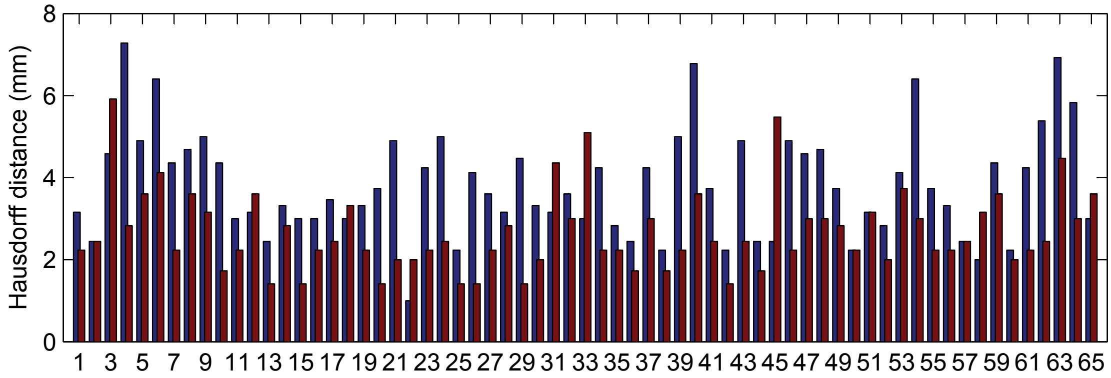
The software to perform the SSWM-guided segmentation is developed in C++ using VTK (Visualization Toolkit) [76] and ITK (Insight Segmentation and Registration Toolkit) [77], and based on one ITK open source project, Deformable Simplex Mesh Model [78]. Our SSWM segmentation software provides the GUI for inspecting the input MR image and the segmentation results in sagittal, coro-



(a) overlap ratio(%). Before deformation: 79.2±9.7/After deformation: 97.3±4.2



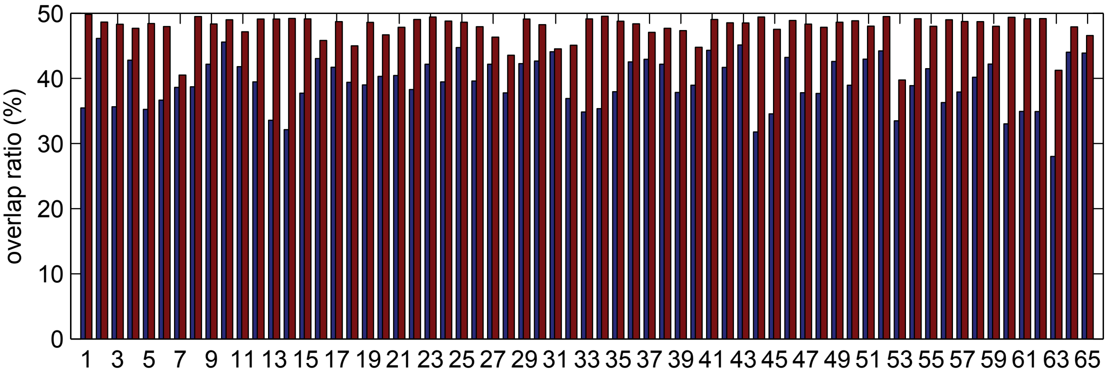
(b) average surface distance(mm). Before deformation: 1.05±0.45/After deformation: 0.21±0.23



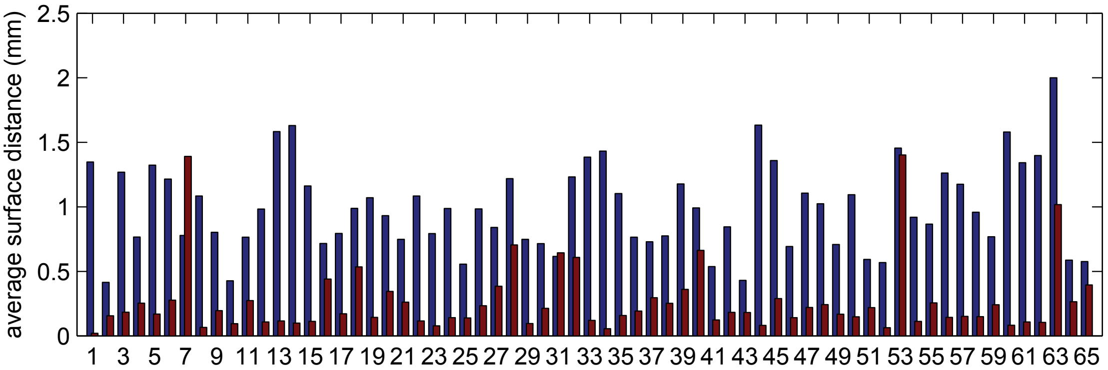
(c) Hausdorff distance(mm). Before deformation: 3.82±1.30/After deformation: 2.68±0.97

Fig. 4.13: Segmentation results of 65 left putamen. Bars in blue illustrate the measure at initialization and in red after deformation

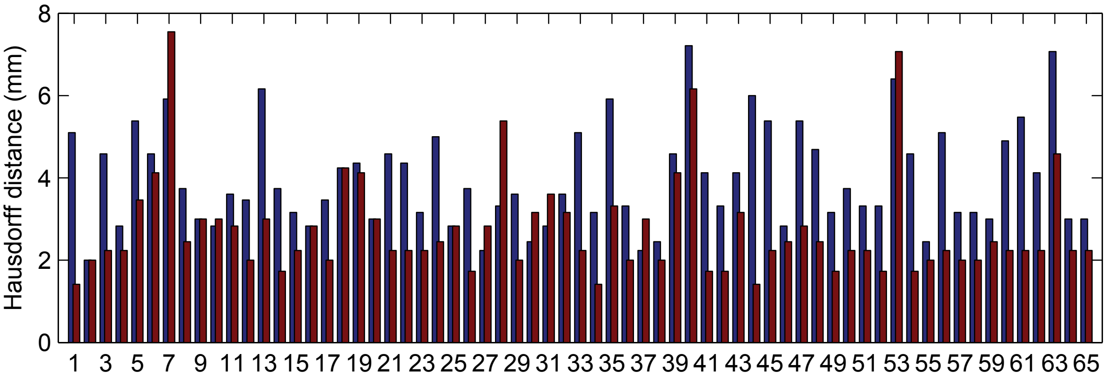
nal and axial directions slide by slide. Screenshots of this software are shown in Fig. 4.15. In Fig. 4.15(a), the prior model is just loaded and initially set to the mean shape. The bottom left window of the GUI shows the current shape (red



(a) overlap ratio(%). Before deformation: 78.9±7.8/After deformation: 95.6±4.2



(b) average surface distance(mm). Before deformation: 0.99±0.34/After deformation: 0.27±0.27



(c) Hausdorff distance(mm). Before deformation: 3.98±1.22/After deformation: 2.75±1.23

Fig. 4.14: Segmentation results of 65 right putamen. Bars in blue illustrate the measure at initialization and in red after deformation

mesh) of the model. There are also three planes perpendicular to each other in this window. These planes cut the mesh and volumetric image. The intersections are shown in the other three windows (the red contours are the intersections of the mesh cut by these planes). We can see, in the example shown in this figure, the initial shape is quite far away from the target shape. Fig. 4.15(b) shows the segmentation results after optimization, in which the model has been fitted to the boundary of caudate nucleus.

Because the software project is managed by Cmake (Cross Platform Make) [79] and the GUI is written in FLTK (Fast Light Toolkit) [80] (with cross-platform support), the program can be compiled and run on different platforms (windows, linux/unix and OSX).

Another key issue in implementing the algorithm is the data-structure used for subdivision wavelet transform. From Fig. 3.1, we see that in Catmull-Clark subdivision, every time, one facet in sub-mesh is divided into four facets in super-mesh. Therefore, the quadtree is the key data structure. Because the classic implementation of quadtree based on pointers requires complex navigation between nodes within one tree or between tress to find the 1-ring/neighbourhood and ancestors/children, we adopted a new technique of quadtree [81, 82] which provides efficient traversal methods as well as efficient storage. It is based on an efficient indexing scheme as a linear (pointerless) quadtree data structure. Such a quadtree is stored using a one-dimensional array of nodes. This indexing scheme has the property that the navigation between any pair of nodes can be computed in constant time. Moreover, the navigation across multiple quadtrees can be achieved at the same cost as well. The adoption of this data-structured boosts the efficiency of our software greatly.

4.5 Conclusion

In this Chapter, a framework of model-guided segmentation using the statistical surface wavelet model has been proposed. Experiment of using this method to segment caudate nuclear on 36 normal cases and 29 cases with schizophrenia has been conducted. After the results being validated with manual segmentations using three segmentation error metrics, the outcome shows that our segmentation method is robust, computationally efficient and achieves a high degree of segmentation accuracy. Moreover, because the multi-variant optimization can be divided into many uni-variant optimization and solved by simple equal-distance sampling, our segmentation method is quiet easy to implement.

In next chapter, based on the segmentation acquired in this chapter, we will compute the comparative group statistics in normal and diseased group respectively. The group mean difference hypothesis will be tested as well to find regions with significant shape change between the 2 groups.

(a) The prior model just loaded

(b) The segmentation results

Fig. 4.15: The software for statistical surface wavelet model (SSWM) guided segmentation.

Chapter 5

Comparative Shape Analysis

In the previous chapter, a method of model-guided segmentation using SSWM has been proposed and tested on 65 caudate nucleus from both schizophrenia patients and healthy controls. In this chapter, a comparative shape analysis between the two shape groups will be conducted to find out whether there is significant shape difference between the two groups, or in other words, whether the shape of the caudate nucleus is affected by schizophrenia.

5.1 Selection of the Datasets

Generally, the 65 MR scans we segmented in the previous chapter can be classified into two groups, 38 schizophrenia patients and 27 normal controls, which are denoted by SP_{all} and NC_{all}, respectively. However, within these two groups, in fact, subjects have other different properties, such as gender, handedness and race. Because of the relatively unknown contributions of gender, handedness and race to the shape analysis, the datasets have to be controlled by selecting subjects based on only one particular value for these properties [83]. Here, we choose "male", "right-handed" and "Chinese" to have a bigger size of the datasets, since these are the most common value observed. According to this selection criteria, we also choose two sub-groups from SP_{all} and NC_{all}: (1) 17 schizophrenics represented by SP_{rhcm}, and (2) 8 normal controls represented by NC_{rhcm}, all of whom are right-handed Chinese male. In the next section, comparative statistical shape analysis of both left and right caudate nucleus is firstly conducted between SP_{all} and NC_{all}, and then between SP_{rhcm} and NC_{rhcm}.

5.2 The Method and Results

Quantitative statistical morphologic analysis of individual brain structures is firstly based on volumetric measurements. These volume changes are only intuitive features as they might explain atrophy or dilation due to illness. However, on the other hand, structural changes at specific locations are not sufficiently reflected in the changes of volume measurements. Shape analysis has thus become of increasing interest to the neuroimaging community due to its potential to precisely locate morphological changes between healthy and pathological structures [13, 83–86]. Our analysis belongs to this category as well. The input of our analysis is the re-meshed surfaces with correspondences, which are generated by processing the binary segmented caudate nucleus using the mesh preparation method described in Section 3.2.

Firstly, the analysis is conducted within two groups, SP_{all} and NC_{all}, on left and right caudate nucleus, respectively. The group mean shape is computed for each group by averaging the 3D coordinates of corresponding surface points across the group, and the results are shown in Fig. 5.1. In order to illustrate the shape difference between the mean shapes of SP_{all} and NC_{all}, at every sample point on the surface, we compute a distance vector representing the direction of local Euclidean surface distance from the mean shape of NC_{all} to the mean shape of SP_{all}. In Fig. 5.2, these distance vectors on every points on the mesh are shown as a distance map. From this figure, we can see that the most different parts between the two mean shapes are the anterior head, posterior tail and superior body (for the anatomical orientation of caudate nucleus, please refer to Fig. 4.1).

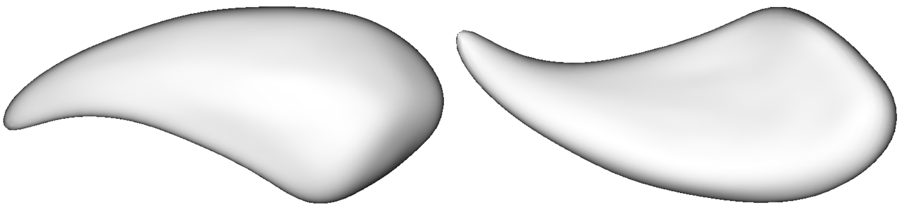
Within every group, we can also show the shape variance at different locations by computing the covariance matrix of the 3D coordinates at every surface points [65]. The eigenvectors of this covariance matrix (3 × 3) are the eigen local variation directions. The eigenvectors of these covariance matrix can be visualized as ellipsoids as shown in Fig. 5.3, in which the three axes of ellipsoid represent three eigenvectors respectively. We can see that, for both left and right caudate nucleus, the variability is reduced in the body section compared to the anterior head, and especially the posterior tail part.

The similar comparisons can be conducted between SP_{rhcm} and NC_{rhcm} as well. Fig. 5.4 shows the mean shapes of group SP_{rhcm} and NC_{rhcm}. Fig. 5.5 shows the distance map from the mean shape of NC_{rhcm} to the mean shape of SP_{rhcm}. Fig. 5.6 shows the covariance ellipsoids within group SP_{rhcm} and NC_{rhcm}. From these figures, similar results are found: anterior head and posterior tail are the parts where there are more shape variance both within each group and between two groups.

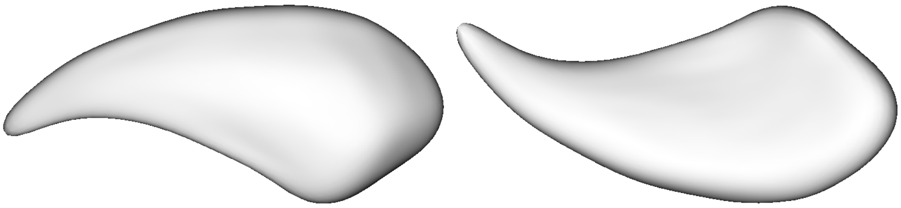
The next step in the comparative shape analysis is the testing for differences between groups at every surface location. This can be done in 2 main fashions. First method is by analyzing the magnitude of the local surface distance vector to a template. For this option, a template needs to be first selected, usually this is the common mean of the 2 groups [85]. The main disadvantage of this method is the need to select a template, which introduces an additional bias into the statistical analysis. The other method is by analyzing the spatial location of each point, in which no template is necessary. Here, we choose the second method and use a commonly accepted multivariate statistics, Hotelling T^2 two sample difference metric [88] as a measurement of how 2 groups locally differ from each other. The Hotelling T^2 is defined as:

$$t^2 = \frac{n_x n_y}{n_x + n_y} (\overline{\mathbf{x}} - \overline{\mathbf{y}})' \mathbf{W}^{-1} (\overline{\mathbf{x}} - \overline{\mathbf{y}}) \tag{5.1}$$

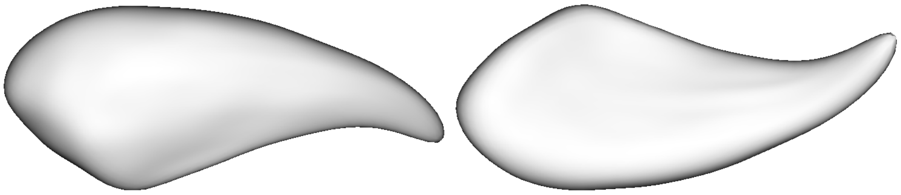
where n_x and n_y are the number of samples in the two groups, $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are the



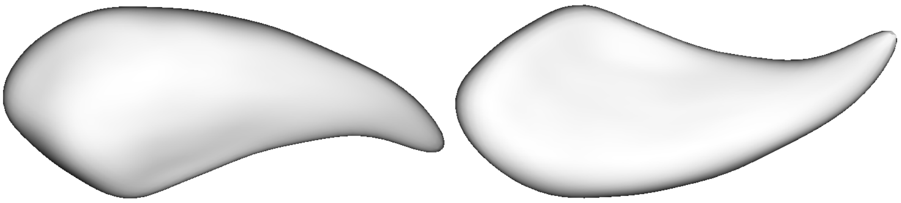
(a) mean shape of the left caudate nucleus in NC_{all} (shown in 2 views)



(b) mean shape of the left caudate nucleus in SP_{all} (shown in 2 views)



(c) mean shape of the right caudate nucleus in NC_{all} (shown in 2 views)



(d) mean shape of the right caudate nucleus in SP_{all} (shown in 2 views)

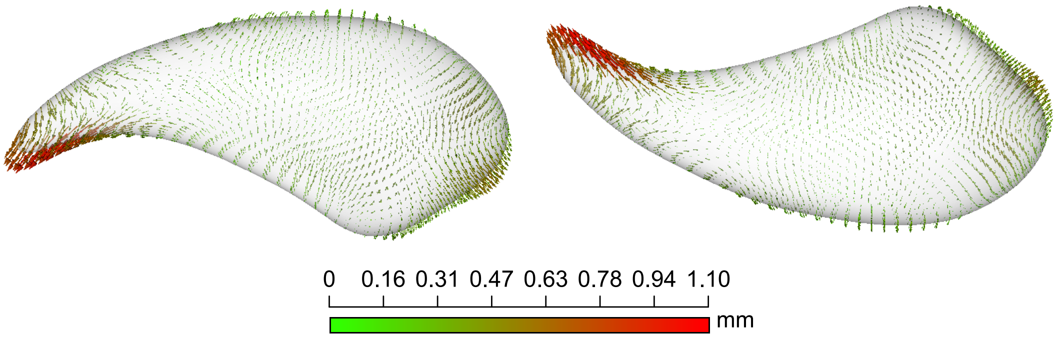
Fig. 5.1: The mean shape of the left and right caudate nucleus in NC_{all} and SP_{all}. (this figure and other figures in this chapter are drawn by software KWMeshVisu [87])

means of the two groups, and

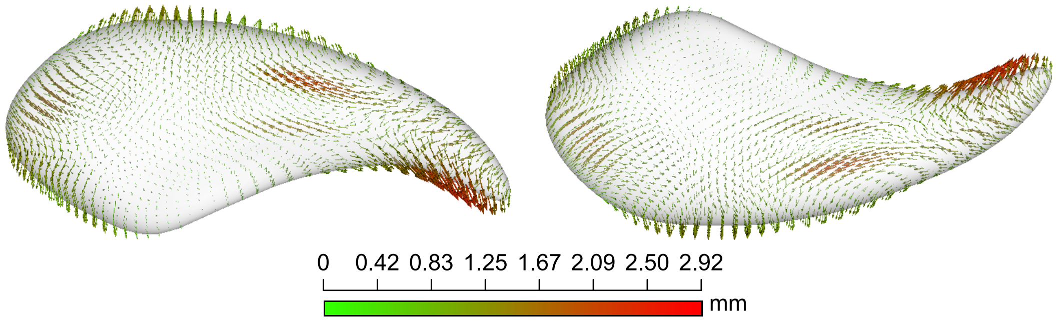
$$\mathbf{W} = \frac{\sum_{i=1}^{n_x}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})' + \sum_{i=1}^{n_y}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})'}{n_x + n_y - 2} \tag{5.2}$$

is the unbiased pooled covariance matrix which is less sensitive to group differences of the covariance matrix and the number of samples.

Using the Hotelling T^2 statistic, the group mean difference hypothesis is tested by using the UNC Statistical Shape Analysis Tools [65] at significance level (α) of



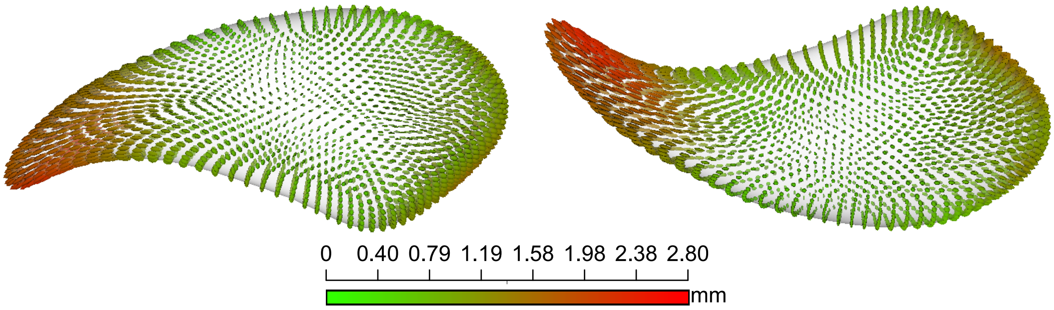
(a) left caudate nucleus (shown in 2 views)



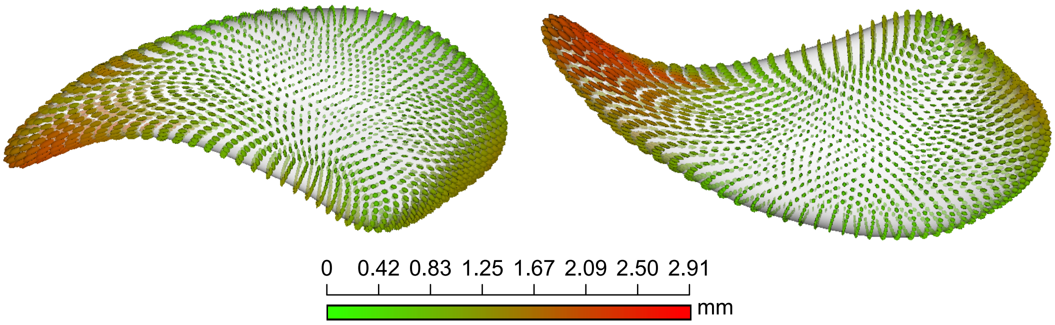
(b) right caudate nucleus (shown in 2 views)

Fig. 5.2: Surface distance between the mean shape in NC_{all} and the mean shape in SP_{all}. The vectors start at the mean shape of NC_{all} and point to the mean shape of SP_{all}.

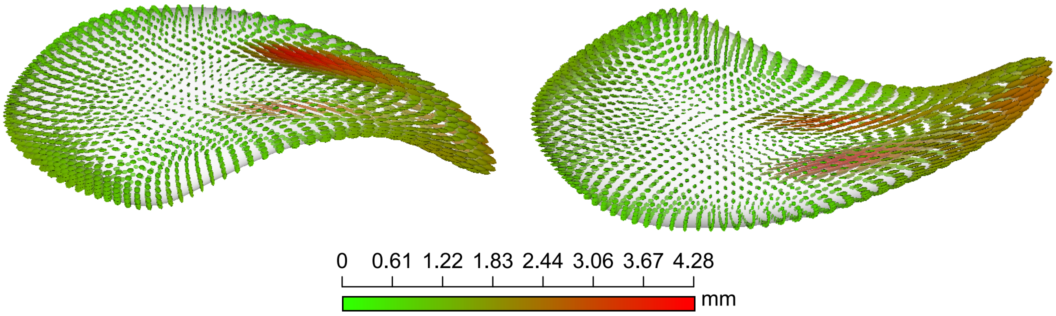
0.05. Our null hypothesis is that the distribution of the locations of each surface point is the same for every subject regardless of the group. The result of this test is a significance map that represents the significance of the local statistical shape difference along the surface as shown in Fig. 5.7(a) and Fig. 5.7(c), where the smaller the p-value the clearer the shape difference is suggested. Here, the significance threshold is selected as 0.05. In these two figures, in some areas, significance group mean shape differences are clearly suggested on both left and right caudate nucleus. However, because of the multiple comparison problem [89], Fig. 5.7(a) and Fig. 5.7(c) are just raw p-value map, and they are overly optimistic estimate of the real significance. Therefore, a correction method based the permutation tests [86, 90] is adopted to correct the above raw p-value map and to control the false positive error. The corrected p-value maps are shown in Fig. 5.7(b) and



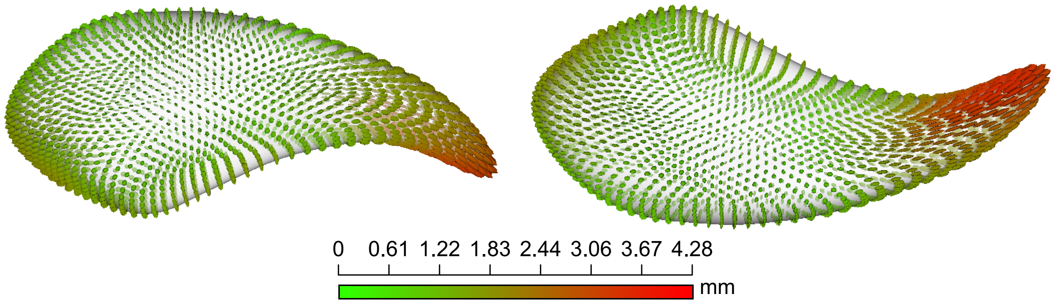
(a) Covariance ellipsoid of left caudate nucleus in NC_{all} (shown in 2 views)



(b) Covariance ellipsoid of left caudate nucleus in SP_{all} (shown in 2 views)

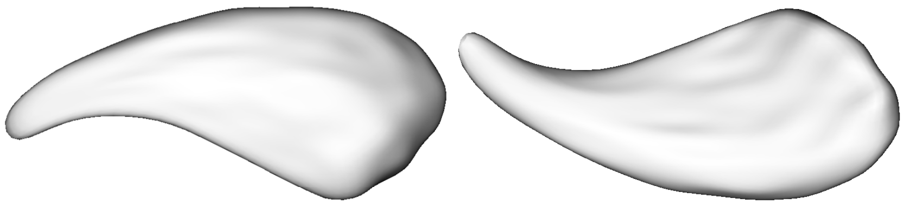


(c) Covariance ellipsoid of right caudate nucleus in NC_{all} (shown in 2 views)

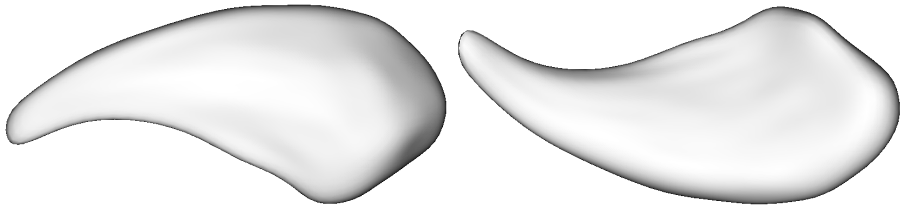


(d) Covariance ellipsoid of right caudate nucleus in SP_{all} (shown in 2 views)

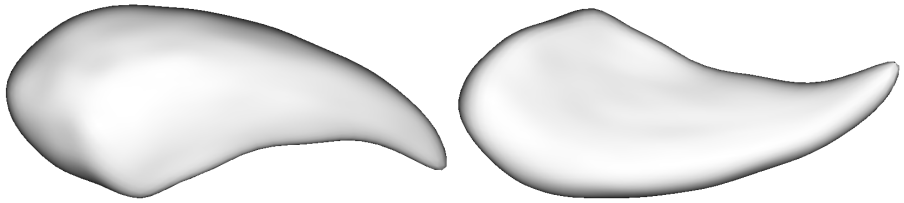
Fig. 5.3: Covariance ellipsoid of left and right caudate nucleus in NC_{all} and SP_{all}.



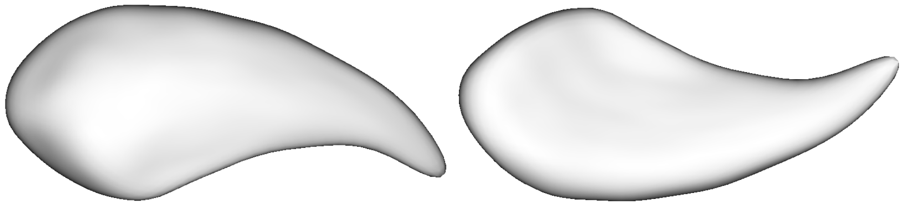
(a) mean shape of the left caudate nucleus in NC_{rhcm} (shown in 2 views)



(b) mean shape of the left caudate nucleus in SP_{rhcm} (shown in 2 views)



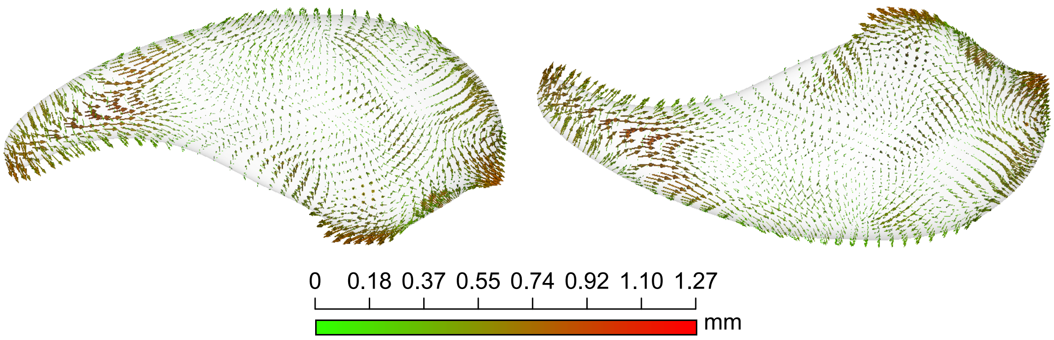
(c) mean shape of the right caudate nucleus in NC_{rhcm} (shown in 2 views)



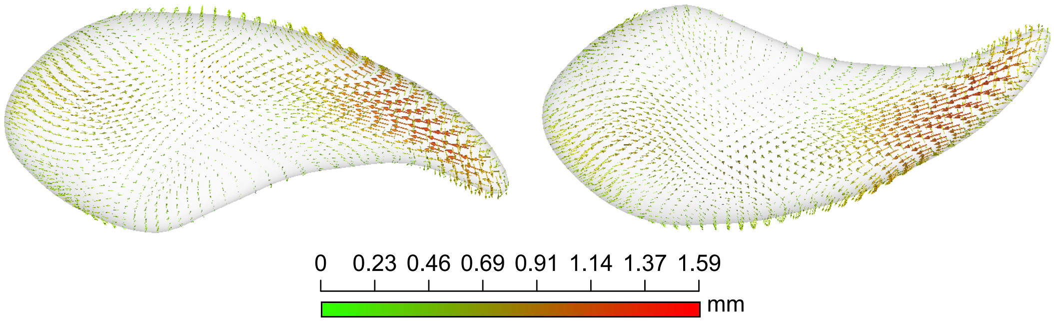
(d) mean shape of the right caudate nucleus in SP_{rhcm} (shown in 2 views)

Fig. 5.4: The mean shape of the left and right caudate nucleus in NC_{rhcm} and SP_{rhcm}.

Fig. 5.7(d), in which significance group mean difference area shrinks compared with the raw p-value map. However, on both left and right caudate nucleus at the anterior head, significant area is still visible. Based on these results, significant group shape difference between NC_{all} and SP_{all} is clearly suggested. In order to exclude the unknown affects of gender, race and handedness to this shape analysis, the same hypothesis testing is also conducted between NC_{rhcm} and SP_{rhcm}. The raw p-value map is shown in Fig. 5.8(a) and Fig. 5.8(c), in which significant areas are detected on both left and right caudate nucleus. But in the corrected p-



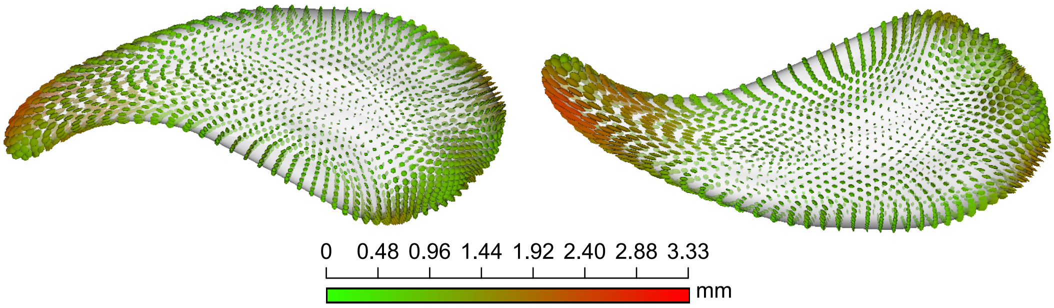
(a) left caudate nucleus (shown in 2 views)



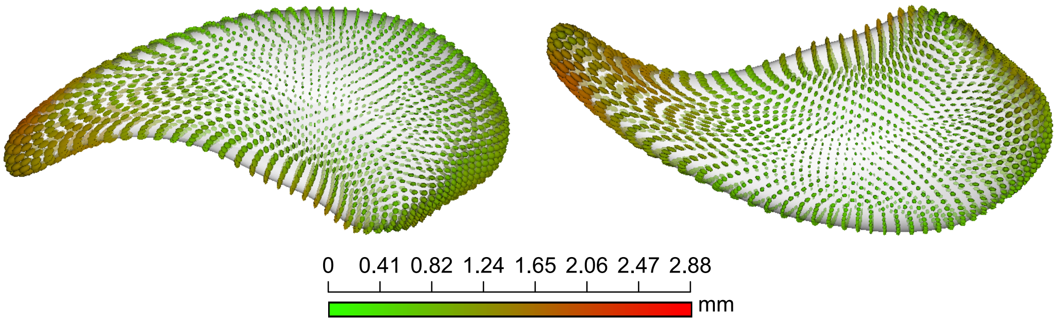
(b) right caudate nucleus (shown in 2 views)

Fig. 5.5: Surface distance between the mean shape in NC_{rhcm} and the mean shape in SP_{rhcm}. The vectors start at the mean shape of NC_{rhcm} and point to the mean shape of SP_{rhcm}.

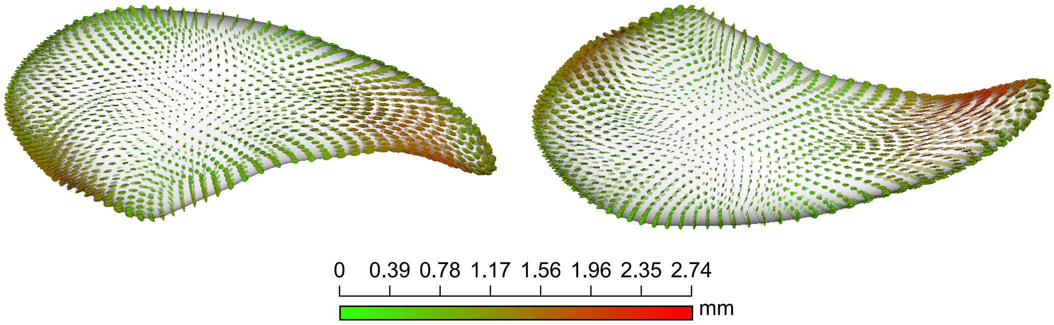
value maps as shown in Fig. 5.8(b) and Fig. 5.8(d), no significant area is detected. This results indicates that significant group shape difference between NC_{rhcm} and SP_{rhcm} is not clearly suggested. Considering the relative insufficient subjects in this analysis (only 17 in SP and 8 in NC), a further study based on more datasets is necessary. We plan to do this research in the future.



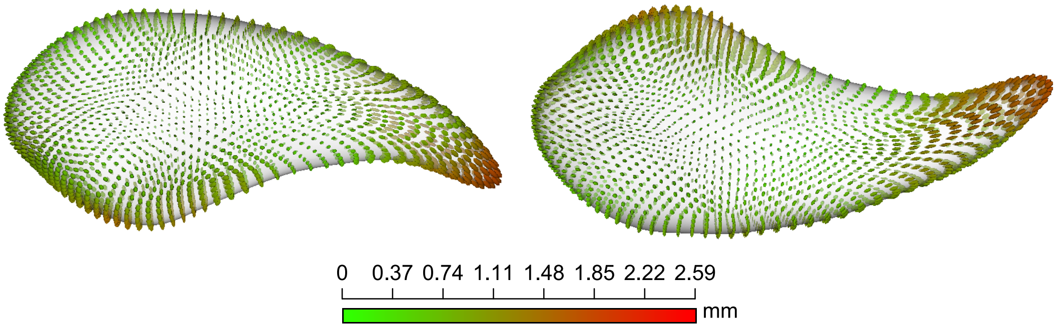
(a) Covariance ellipsoid of left caudate nucleus in NC_{rhcm} (shown in 2 views)



(b) Covariance ellipsoid of left caudate nucleus in SP_{rhcm} (shown in 2 views)

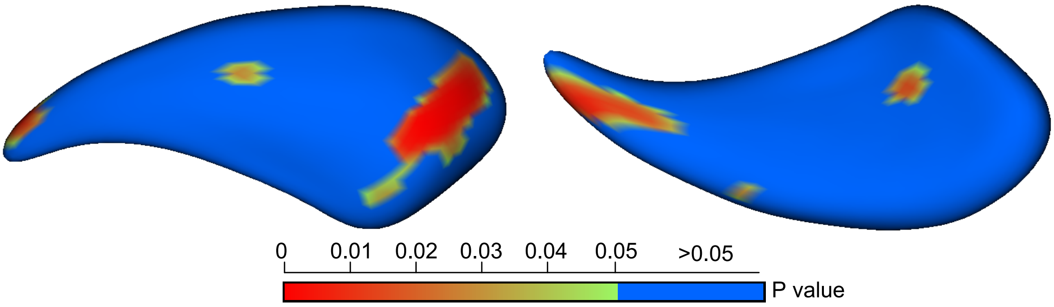


(c) Covariance ellipsoid of right caudate nucleus in NC_{rhcm} (shown in 2 views)

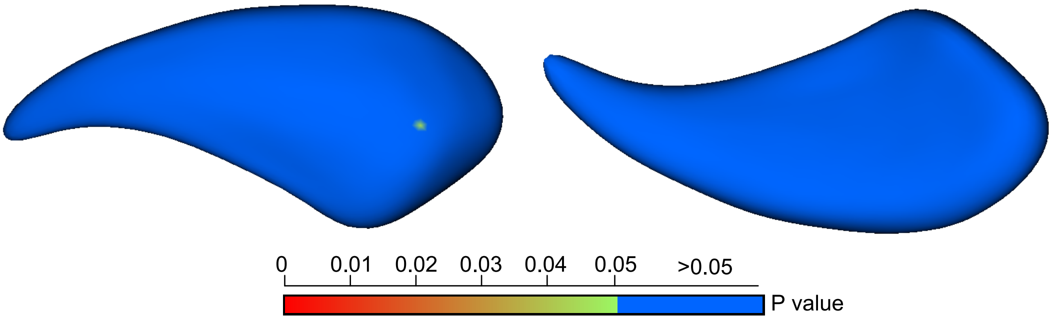


(d) Covariance ellipsoid of right caudate nucleus in SP_{rhcm} (shown in 2 views)

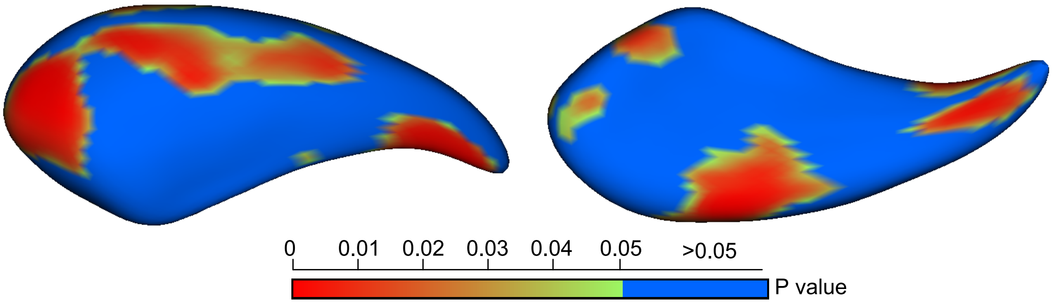
Fig. 5.6: Covariance ellipsoid of left and right caudate nucleus in NC_{rhcm} and SP_{rhcm}.



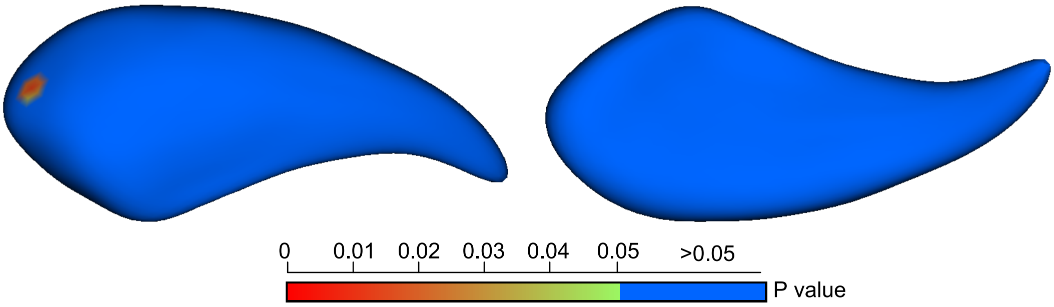
(a) The raw (optimistic) significance map of the left caudate nucleus (shown in 2 views)



(b) The corrected significance map of the left caudate nucleus (shown in 2 views)

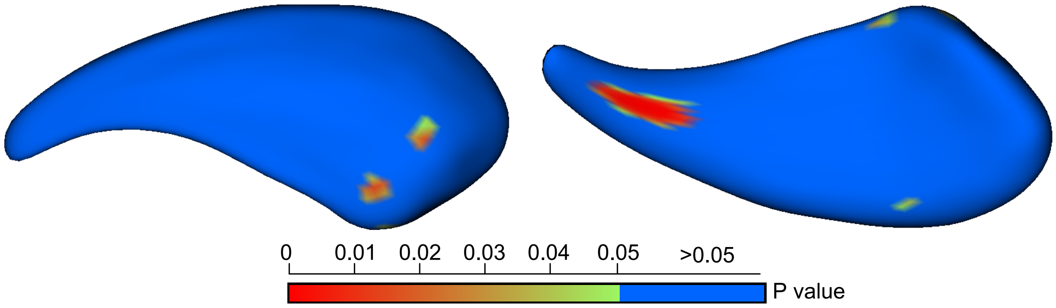


(c) The raw (optimistic) significance map of the right caudate nucleus (shown in 2 views)

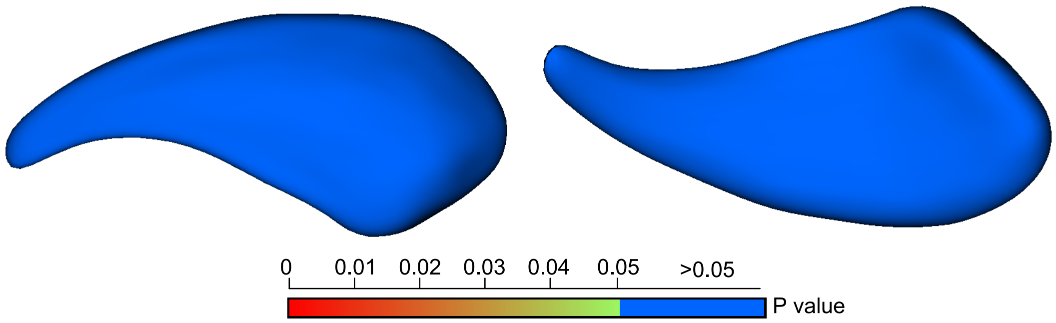


(d) The corrected significance map of the right caudate nucleus (shown in 2 views)

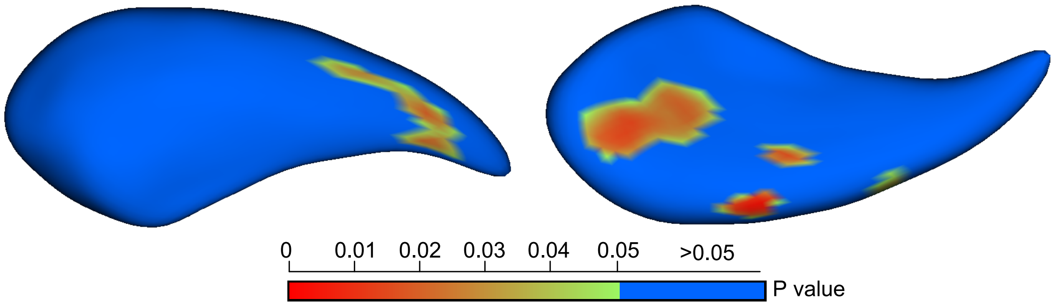
Fig. 5.7: Group mean shape difference testing between NC_{all} and SP_{all}.



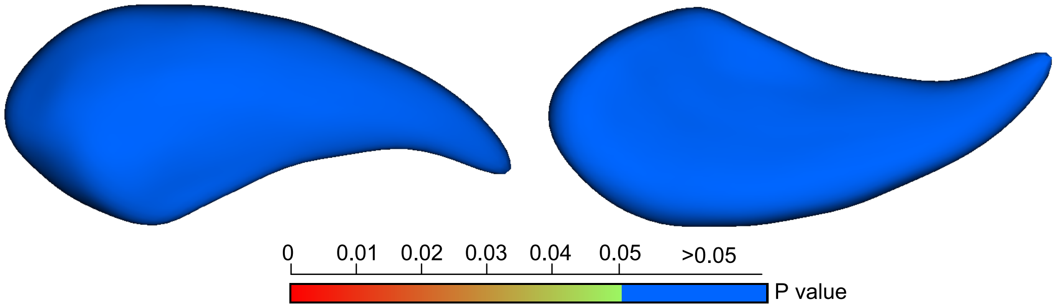
(a) The raw (optimistic) significance map of the left caudate nucleus (shown in 2 views)



(b) The corrected significance map of the left caudate nucleus (shown in 2 views)



(c) The raw (optimistic) significance map of the right caudate nucleus (shown in 2 views)



(d) The corrected significance map of the right caudate nucleus (shown in 2 views)

Fig. 5.8: Group mean shape difference testing between NC_{rhcm} and SP_{rhcm}.

Chapter 6

Conclusion and Future Work

This research investigated the possibility of utilizing the advantage of the wavelets analysis in statistical shape analysis of 3D surfaces. Using a recently developed technique, subdivision surface wavelet, a new statistical shape model called SSWM (Statistical Surface Wavelet Model) was proposed. Due to the ability of the adopted wavelet scheme to perform biorthogonal wavelet analysis directly on the surface mesh with Catmull-Clark subdivision mesh connectivity, this model avoids the difficulties caused by the explicit surface parameterization and extends the wavelet-based shape analysis and statistical shape model building from 2D to 3D successfully. Hence, besides the multiscale statistical shape analysis which other methods, for example Fourier and Spherical Harmonics, can perform as well, it has an advantage of being able to conduct a spatially localized multiscale analysis. This means it can focus the shape analysis on a certain area of the surface to a desired scale while the other part of the surface remains unchanged. Such a statistical shape analysis was performed on 18 samples of caudate nucleus from the IBSR [8] datasets. The analysis results clearly showed the different shape variations in different scales at different spatial locations. This capability makes it possible to find out which part of the organ will most likely be affected by the disease.

A framework of using SSWM for automatic model-guided segmentation of deep cerebral structures was also developed, in which SSWM was used as a shape prior. In model-guided segmentation, the model-image fitting is actually a multi-variable global optimization problem usually of the order of several tens to a few hundreds free parameters. Hence, no analytical solution to this problem is available. Possible optimization algorithms are usually deterministic techniques like conjugate gradient or stochastic methods such as genetic algorithms. By using SSWM's multiscale and spatially localized description of the shape, the model-image fitting problem was successfully converted from one multi-variable optimization problem to many single-variable optimization problems. In this divide-and-conquer manner, the optimization problem can be solved much more efficiently and reliably. We have tested our method on 65 MR images to segment caudate nucleus and putamen using the SSWM priors trained from IBSR datasets. The results were very encouraging. In all 65 samples, the proposed method successfully segmented the right and left caudate nucleus. The method also successfully segmented most of the left and right putamens, except some cases in which the boundary of putamen is partially missing (Scenario A, in Section 4.3). The computation time for this automatic model-guided segmentation is about 3 minutes on a P4 2.4G Windows XP system. By using the Talairach coordinate, our model integrates the shape information together with the similarity transform information and results in a more robust segmentation method which is not sensitive to the model initialization. We also found that our method can separate caudate and putamen correctly and doesn't cause leakage as the other model-free methods do. This is the clear evidence of the successful utilization of prior anatomical knowledge. In the future, we will increase the size of the training set to get a more comprehensive and accurate shape prior and test our method on more brain structures.

Currently the proposed SSWM can only handle genus-0 closed surface. In fact, this constrain is not imposed by the surface wavelet scheme [48] which can process mesh surface with any genus as long as it is closed and has the required Catmull-Clark subdivision mesh connectivity. The reason why this framework can't be extended to higher genus is because of the lack of an appropriate correspondence-finding and re-meshing algorithm for the surface with a higher genus. The correspondence-finding and re-meshing algorithm introduced in Section 3.2 is only applicable to closed genus-0 surface. If this problem can be solved, the proposed model can be extended without any change. Regarding the open surface, the most convenient way is to patch the open area to make the surface closed. Since the wavelets are spatially localized, the patch will only result in a number of useless wavelets coefficients (which are corresponding to the patched area). In the subsequent steps, these coefficients can just be omitted. However, this patching method can not be used for other models, such as SPHARM [5] or ASM [2], because without the spatial-localization property, the patch will affect other coefficients of the model.

After the segmentation of caudate nucleus from 65 MR scans, a comparative shape analysis between schizophrenia patient and healthy control was performed in Chapter 5. Firstly, in the statistical group mean difference hypothesis testing between schizophrenia and healthy controls regardless of gender, race and handedness, significant shape difference between the two groups was clearly suggested. In order to exclude the unknown affects of gender, race and handedness to the shape analysis, the same hypothesis testing was also conducted on two sub-groups which only consists of right handed Chinese male. However, in this test, no significant shape difference between the two groups was clearly suggested. Considering the relative insufficient subjects in this analysis (only 17 schizophrenia patients and 8 healthy controls), a further study based on more datasets is necessary and we plan to do this research in the future.

In some other statistical shape analysis work on hippocampus [13, 83, 85], significant shape difference between normal and schizophrenic was reported. Therefore, because of the existence of this shape difference, in their work, a classifier was finally trained to help the diagnosis of this disease. In the future study of caudate nucleus, if shape difference is detected, we also plan to develop a shape classifier based on our SSWM to utilize the spatial localization property of our model. It is known that feature selection is always a problem of classifier construction and training in pattern recognition. In theory, having more coefficients in the shape descriptor as input to the classifier can only improve or, at least, not change the performance of the classifier. However, in practice, only some of the coefficients in the shape descriptor are critical to the classification and the others are less informative. Therefore, mis-estimation of the less informative coefficients can actually degrade the performance of the classifier. This trend of actual losses of accuracy resulting from additional input coefficients is known as the "peaking effect" or "Hughes phenomenon" [91]. Thus, it is often helpful to select a subset of the most useful coefficients from the shape descriptors. These selected coefficients should describe the surface shape at spatial locations and scales where the most significant shape difference between healthy controls and patients occurs. In order to make such a "double" selection possible, the shape model used for classification should be selective not only to scale but also to spatial locations. PDM is only selective to the spatial location on the surface, but not to the scale. SPHARM is multiscale, but because of its global supported basis function, it is only selective to the scale. In contrast to the above 2 models, SSWM provides selectivity both in scale and spatial location for feature selection. Therefore, when used as the shape descriptor, SSWM can potentially improve the classification accuracy.

Bibliography

[1] I. Dryden and K. Mardia, *Statistical shape analysis.* Chichester: Wiley, 1998.

[2] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[3] L. H. Staib and J. S. Duncan, "Model-based deformable surface finding for medical images," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 720–731, Oct. 1996.

[4] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 828–839, Oct. 1999.

[5] G. Szekely, A. Kelemen, C. Brechbuhler, and G. Gerig, "Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible fourier contour and surface models," *Medical Image Analysis*, vol. 1, no. 1, pp. 19–34, 1996.

[6] M. Leventon, W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proceedings of CVPR'2000: Computer Vision and Pattern Recognition*, 2000, pp. 316–323.

[7] S. M. Pizer, D. S. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 851–865, 1996.

[8] The Internet Brain Segmentation Repository (IBSR). [Online]. Available: http://www.cma.mgh.harvard.edu/ibsr/

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Training models of shape from sets of examples," in *Proceedings of British Machine Vision Conference.* Springer-Verlag, 1992, pp. 9–18.

[10] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Med. Imag.*, vol. 14, no. 11, pp. 1061–1075, Nov. 1992.

[11] C. Davatzikos, X. Tao, and D. Shen, “Hierarchical active shape models, using the wavelet transform,” *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 414–423, Mar. 2003.

[12] C. Davatzikos, M. Vaillant, S. Resnick, J. L. Prince, S. Letovsky, and R. N. Bryan, “A computerized method for morphological analysis of the corpus callosum,” *Journal of Computer Assisted Tomography*, vol. 20, no. 1, pp. 88–97, 1996.

[13] P. Golland, W. Grimson, M. Shenton, and R. Kikinis, “Detection and analysis of statistical differences in anatomical shape,” *Medical Image Analysis*, vol. 9, no. 1, pp. 69–86, 2005.

[14] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd ed).* New York: Wiley, 2000.

[15] V. N. Vapnik, *The Nature of Statistical Learning Theory.* Springer, 1995.

[16] ——, *Statistical Learning Theory.* Wiley, 1998.

[17] L. Holly and K. Foley, “Intraoperative spinal navigation,” *Spine*, vol. 28, no. 15, pp. 554–561, 2003.

[18] B.O’Sullivan and J. Shah., “New TNM staging criteria for hear and neck tumors,” *Semin. Surf. Oncol.*, vol. 21, no. 1, pp. 30–42, 2003.

[19] S. M. Haney, P. M. Thompson, T. F. Cloughesy, J. R. Alger, and A. W. Toga, “Tracking tumor growth rates in patients with malignant gliomas: A test of two algorithms,” *American Journal of Neuroradiology*, vol. 22, no. 1, pp. 73–82, Jan. 2001.

[20] J. Chassery and C. Garbay, “An interactive segmentation method based on a contextual color and shape criterion,” *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 794–800, 1984.

[21] W. Zhao, T. Young, and M. Ginsberg, “Registration and three-dimensional reconstruction of autoradiographic images by the disparity analysis method,” *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 782–791, 1993.

[22] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[23] T. McInerney and D. Terzopoulos, “Deformable models in medical image analysis: A survey.” *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.

[24] G. C.-H. Chang and C.-C. J. Kuo, “Wavelet descriptor of planar curves: theory and applications,” *IEEE Trans. Image Processing*, vol. 5, no. 1, pp. 56–70, Jan. 1996.

[25] W. Sweldens, "The lifting scheme: a custom-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis*, vol. 3, no. 2, pp. 186–200, 1996.

[26] F. Bookstein, "Landmark methods for forms without landmarks: Localizing group differences in outline shape," in *Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996.

[27] ——, "Shape and the information in medical images: A decade of the morphometric synthesis," in *Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996.

[28] F. L. Bookstein, "Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape," *Medical Image Analysis*, vol. 1, no. 3, pp. 225–243, 1997.

[29] T. Cootes and C. Taylor, "Active shape models - 'Smart Snakes'," in *British Mach. Vision Conf.* Springer-Verlag, 1992, pp. 266–275.

[30] A. Rangarajan, H. Chui, and F. Bookstein, "The soft assign procrustes matching algorithm," in *International Conference on Information Processing in Medical Imaging.* LNCS 1230, 1997, pp. 29–42.

[31] J. Canny, "A computational approach to edge detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, November 1986.

[32] A. Blake and M. Isard, *Active Contours.* Springer, 1998.

[33] H. Delingette, G. Subsol, S. Cotin, and J. Pignon, "A craniofacial surgery simulation testbed," in *Visualization in Biomedical Computing (VBC'94)*, oct. 1994, pp. 607–618.

[34] S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[35] R. Whitaker, "Algorithms for implicit deformable models," in *Proc. 5th Int. Conf Comp. Vision*, June 1995, pp. 822–827.

[36] R. Haralick and L. Shapiro, *Computer and Robot Vision.* Addison-Wesley Publishing Company, 1992, vol. 1.

[37] J. Yang, L. H. Staib, and J. S. Duncan, "Neighbor-constained segmentation with level set based 3-D deformable models," *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 940–948, 2004.

[38] R. Bajcsy and F. Solina, "Three dimensional object representation revisited," in *1st International Conference on Computer Vision (ICCV)*, London, June 1987, pp. 231–240.

[39] E. Bardinet, L. D. Cohen, and N. Ayache, "Fitting the 3-d data using superquadrics and free-form deformations," in *12th IAPR International Conference on Pattern Recognition*, 1994, pp. 79–83.

[40] B. C. Vemuri and A. Radisavljevic, "Multiresolution stochastic hybrid shape models with fractal priors," *ACM Transaction on Graphics*, vol. 13, no. 2, pp. 177–207, April 1994.

[41] C. Brechbhler, G. Gerig, and O. Kubler, "Parametrization of closed surfaces for 3D shape description," *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 154–170, Mar. 1995.

[42] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge Univ. Press, 1993.

[43] Spherical Harmonics Generator. [Online]. Available: http://adomas.org/shg/

[44] M. Quicken, C. Brechbühler, J. Hug, H. Blattman, and G. Székely, "Parameterization of closed surfaces for parametric surface description," in *Proceedings of CVPR'2000: Computer Vision and Pattern Recognition*, 2000, pp. 354–360.

[45] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[46] I. Daubechies, *Ten Lectures on Wavelets*, ser. CBMS-NSF Regional Conferences Series in Applied Mathematics. Philadelphia PA: SIAM, 1992.

[47] P. Goupillaud, A. Grossman, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," in *Geoexploration*, 1985, pp. 85–102.

[48] M. Bertram, M. A. Duchaineau, B. Hamann, and K. I. Joy, "Generalized B-spline subdivision-surface wavelets for geometry compression," *IEEE Trans. Visual. Comput. Graphics*, vol. 10, no. 3, pp. 326–338, May/June 2004.

[49] P. Schröder and W. Sweldens, "Spherical wavelets: efficiently representing functions on the sphere," *Computer Graphics*, vol. 29, no. Annual Conference Series, pp. 161–172, 1995.

[50] Y. Peng, P. E. Grant, Y. Qi, X. Han, F. Segonne, R. Pienaar, E. Busa, J. Pacheco, N. M. R. L. Buckner, P. Golland, and B. Fischl, "Cortical surface shape analysis based on spherical wavelets," *IEEE Trans. Med. Imag.*, vol. 26, no. 4, pp. 582–597, Apr. 2007.

[51] P. Yu, X. Han, F. Segonne, R. Pienaar, R. L. Buckner, P. Golland, P. E. Grant, and B. Fischl, “Cortical surface shape analysis based on spherical wavelet transformation,” in *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis*, New York City, June 2006, p. 60.

[52] D. Nain, A. B. S. Haker, and A. R. Tannenbaum, “Multiscale 3-D shape representation and segmentation using spherical wavelets,” *IEEE Trans. Med. Imag.*, vol. 26, no. 4, pp. 598–618, April 2007.

[53] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum, “Shape-driven 3D segmentation using spherical wavelets,” in *Proceedings of MICCAI’2006: Medical Image Computing and Computer-Assisted Intervention*, Oct. 2006, pp. 66–74.

[54] D. Nain, S. Haker, A. Bobick, and A. R. Tannenbaum, “Multiscale 3d shape analysis using spherical wavelets,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI).* Palm Springs, CA, USA: Springer, October 2005, pp. 459–467.

[55] J. M. Lounsbery, “Multiresolution analysis for surfaces of arbitrary topological type,” Ph.D. dissertation, Department of Computer Science and Engineering, University of Washington, Sept. 1994.

[56] J. M. Lounsbery, T. DeRose, and J. Warren, “Multiresolution analysis for surfaces of arbitrary topological type,” *ACM Trans. Graphics*, vol. 16, no. 1, pp. 34–73, 1997.

[57] E. Catmull and J. Clark, “Recursively generated b-spline surfaces on arbitrary topological meshes,” *Comput. Aided Des.*, vol. 10, no. 350-355, 1978.

[58] C. Loop, “Smooth subdivision surfaces based on triangles,” Master’s thesis, Univ. of Utah, Salt Lake City, 1987.

[59] P. Besl and N. McKay, “A method for registration of 3-d shapes,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[60] A. Kotche and C. Taylor, “Automatic construction of eigenshape models by direct optimization,” *Medical Image Analysis*, vol. 2, no. 4, pp. 303–314, dec 1998.

[61] R. Davies, C. Twining, T. Cootes, and J. C. Waterton, “A minimum description length approach to statistical shape model,” *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 525–537, may 2002.

[62] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, pp. 465–471, 1978.

[63] Y. Wang, B. Peterson, and L. Staib, "Shape-based 3d surface correspondence using geodesics and local geometry," in *Proceedings of CVPR'2000: Computer Vision and Pattern Recognition*, 2000, pp. 644–651.

[64] M. Styner, K. Rajamani, L. P. Nolte, G. Szekely, C. Talor, and R. H. Davies, "Evaluation of 3D correspondence methods for model building," in *Information Processing in Medical Imaging (IPMI)*, 2003, pp. 63–75.

[65] Shape Analysis Tools (UNC). [Online]. Available: http://www.ia.unc.edu/dev/download/shapeAnalysis/

[66] C. Zahn and R. Roskies, "Fourier descriptors for plane close curves," *IEEE Trans. Computers*, vol. C-21, pp. 269–28, March 1972.

[67] J. Talairach and P. Tournoux, *Co-Planar stereotaxic atlas of the human brain.* Stuttgart, Germany: Thieme, 1988.

[68] W. Nowinski, K. B. Prakash, I. Volkau, A. Ananthasubramaniam, and N. Beauchamp, "Rapid and automatic calculation of the midsagittal plane in magnetic resonance diffusion and perfusion images." *Academic Radiology*, vol. 13, no. 5, pp. 652–663, 2006.

[69] Q. Hu and W. Nowinski, "Radiological symmetry of brain and head images: comparison and applications." *International Journal of Computer Assisted Radiology and Surgery*, vol. 1, no. 2, pp. 75–81, 2006.

[70] I. Volkau, K. B. Prakash, A. Anand, A. Aziz, and W. Nowinski, "Extraction of the midsagittal plane from morphological neuroimages using the Kullback-Leiblers measure." *Medical Image Analysis*, vol. 10, no. 6, pp. 863–874, 2006.

[71] W. Nowinski, G. Qian, K. B. Prakash, Q. Hu, and A. Aziz, "Fast talairach transformation for magnetic resonance neuroimages." *Journal of Computer Assisted Tomography*, vol. 30, no. 4, pp. 629–641, July/August 2006.

[72] Q. Hu, G. Qian, and W. Nowinski, "Fast, accurate and automatic extraction of the modified talairach cortical landmarks from mr images." *Magnetic Resonance in Medicine*, vol. 53, pp. 970–976, 2005.

[73] Caudate nucleus, Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/caudate_nucleus

[74] Valmet. [Online]. Available: http://www.ia.unc.edu/dev/download/valmet/index.htm

[75] Putamen, Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Putamen

[76] The Visualization ToolKit (VTK). [Online]. Available: http://www.vtk.org/

[77] The Insight Segmentation and Registration Toolkit (ITK). [Online]. Available: http://www.itk.org/

[78] "The insighttoolkit applications." [Online]. Available: http://www.itk.org/HTML/Applications.htm

[79] CMake, the cross-platform, open-source make system (Cmake). [Online]. Available: http://www.cmake.org/

[80] The Fast Light Toolkit (FLTK). [Online]. Available: http://www.fltk.org/

[81] L. Balmelli, S. Ayer, and M. Vetterli, "Efficient algorithms for embbeded terrain simplification," in *IEEE International Conference on Image Processing (ICIP)*, Chicago, USA, October 1998.

[82] L. Balmelli, J. Kovacevic, and M. Vetterli, "Quadtrees for embedded surface visualization: Constraints and efficient data structures," in *IEEE International Conference on Image Processing (ICIP)*, Kobe, Japan, October 1999.

[83] L. Shen, J. Ford, F. Makedon, and A. Saykin, "Hippocampal shape analysis surface-based representation and classification," in *SPIE-Medical Imaging*, 2003.

[84] L. Shen, J. Ford, F. Makedon, and A. J. Saykin, "A surface-based approach for classification of 3D neuroanatomic structures," *Intell. Data Anal.*, vol. 8, no. 6, pp. 519–542, 2004.

[85] M. Styner, J. Lieberman, D. Pantazis, and G. Gerig, "Boundary and medial shape analysis of the hippocampus in schizophrenia," *Medical Image Analysis*, vol. 8, no. 3, pp. 197–203, 2004.

[86] D. Pantazis, R. Leahy, T. Nichol, and M. Styner, "Statistical surface-based morphometry using a non-parametric approach," in *Int. Symposium on Biomedical Imaging(ISBI)*, April 2004, pp. 1283–1286.

[87] I. Oguz, G. Gerig, S. Barre, and M. Styner, "Kwmeshvisu: A mesh visualization tool for shape analysis," 2006 MICCAI Open Science Workshop, Oct. 2006. [Online]. Available: http://insight-journal.org/dspace/handle/1926/220

[88] S. GAF, *Multivariate Observations.* John Wiley, 1984.

[89] "The multiple comparison problem, wikipedia." [Online]. Available: http://en.wikipedia.org/wiki/Multiple_comparisons

[90] E. Edgington, *Randomization Tests.* Academic Press, 1995.

[91] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inform. Theory*, vol. 14, no. 1, pp. 55–63, Jan. 1968.

[92] M.Bertram and M. D. B. Hamann, “Bicubic subdivision-surface wavelets for large-scale isosurface representation and visualization,” in *Proceedings of Visualization*, 2000, pp. 389–396.

[93] R. Duda, P. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. John Wiley Sons Inc., 2001.

Appendix A

Generalized B-spline Subdivision-Surface Wavelets

The outline for constructing subdivision surface wavelets is to decompose the subdivision rules into a series of reversible simple local operations and then convert the forward and inverse subdivisions into wavelet transforms. Specifically, the wavelet scheme we adopted in this thesis is defined on the Catmull-Clark subdivision [57] surface, which is a generalization of uniform bicubic B-splines to arbitrary control meshes. A mesh is refined by inserting a new vertex inside every face and on every edge and by connecting these vertices to quadrilaterals. Vertices in a supermesh correspond to a facet (polygon), an edge, or a vertex in the submesh and are denoted by $\mathbf{f}$, $\mathbf{e}$, $\mathbf{v}$, respectively.

For brevity to describe subdivision rules determining new vertex positions, we use the index-free averaging operator $\overline{\mathbf{x}}_{\mathbf{y}}$ introduced in [48, 92] to illustrate, where $\mathbf{x}$ and $\mathbf{y}$ can represent either $\mathbf{f}$, $\mathbf{e}$ or $\mathbf{v}$. This averaging operator returns for every vertex of type $\mathbf{y}$ the arithmetic average of all adjacent vertices of type $\mathbf{x}$. In particular, we use the following notation that is illustrated in Fig. A.1: $\overline{\mathbf{v}}_{\mathbf{f}}$: centroid of each face; $\overline{\mathbf{e}}_{\mathbf{f}}$: centroid of $\mathbf{e}$ vertices of each face; $\overline{\mathbf{v}}_{\mathbf{e}}$: mid-point of each edge; $\overline{\mathbf{f}}_{\mathbf{e}}$: midpoint of both adjacent $\mathbf{f}$ vertices of each edge; $\overline{\mathbf{v}}_{\mathbf{v}}$: centroid of all adjacent $\mathbf{v}$ vertices; $\overline{\mathbf{e}}_{\mathbf{v}}$: centroid of all $\mathbf{e}$ vertices of incident edges; $\overline{\mathbf{f}}_{\mathbf{v}}$: centroid of

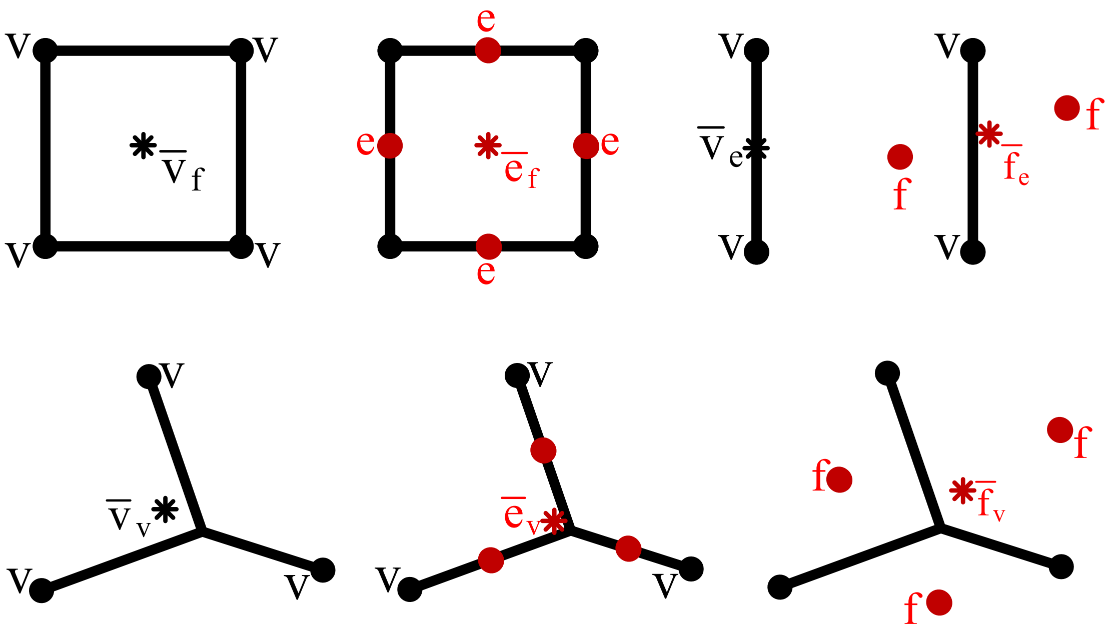


Fig. A.1: The index-free notation for subdivision surface wavelet transform. $\overline{\mathbf{v}}_{\mathbf{f}}$ denotes the centroid of a facet, $\overline{\mathbf{e}}_{\mathbf{f}}$ denotes the centroid of the associated **e** vertices, etc.

all **f** vertices of incident faces.

With the above index-free notation, Catmull-Clark subdivision is defined by the rules:

$$\begin{aligned}
\mathbf{f}' &\longleftarrow \overline{\mathbf{v}}_{\mathbf{f}} \\
\mathbf{e}' &\longleftarrow \tfrac{1}{2}(\overline{\mathbf{v}}_{\mathbf{e}} + \overline{\mathbf{f}'}_{\mathbf{e}}) \\
\mathbf{v}' &\longleftarrow \tfrac{1}{n_{\mathbf{v}}}(\overline{\mathbf{f}'}_{\mathbf{v}} + \overline{\mathbf{v}}_{\mathbf{v}} + (n_{\mathbf{v}} - 2)\mathbf{v})
\end{aligned} \tag{A.1}$$

where $n_{\mathbf{v}}$ is the valence (number of incident edges) of vertex $\mathbf{v}$. Here, the order of vertex modifications is important because the result of an operation may define the input of the subsequent operations.

Similarly, the vertex modification rules fore the generalized B-spline wavelet analysis (forward transform) are defined as [48]:

$$
\begin{aligned}
\mathbf{v} &\longleftarrow \mathbf{v} + \tfrac{1}{4}\overline{\mathbf{f}}_{\mathbf{v}} - \overline{\mathbf{e}}_{\mathbf{v}} \\
\mathbf{e} &\longleftarrow \mathbf{e} - \tfrac{1}{2}\overline{\mathbf{f}}_{\mathbf{e}} \\
\mathbf{f} &\longleftarrow \mathbf{f} + 4\overline{\mathbf{v}}_{\mathbf{f}} - 4\overline{\mathbf{e}}_{\mathbf{f}} \\
\mathbf{e} &\longleftarrow \mathbf{e} - 2\overline{\mathbf{v}}_{\mathbf{e}} \\
\mathbf{v} &\longleftarrow 4\mathbf{v} + \tfrac{9}{16}\overline{\mathbf{f}}_{\mathbf{v}} + 3\overline{\mathbf{e}}_{\mathbf{v}} \\
\mathbf{e} &\longleftarrow 2\mathbf{e} + \tfrac{3}{4}\overline{\mathbf{f}}_{\mathbf{e}}
\end{aligned}
\tag{A.2}
$$

And the vertex modification rules fore the generalized B-spline wavelet synthesis (inverse transform) are defined as [48]:

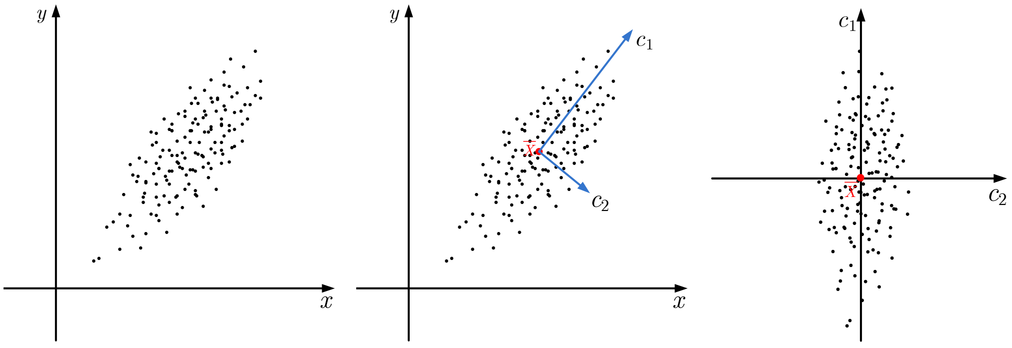
$$
\begin{aligned}
\mathbf{e} &\longleftarrow \tfrac{1}{2}\mathbf{e} - \tfrac{3}{8}\overline{\mathbf{f}}_{\mathbf{e}} \\
\mathbf{v} &\longleftarrow \tfrac{1}{4}\mathbf{v} + \tfrac{9}{64}\overline{\mathbf{f}}_{\mathbf{v}} - \tfrac{3}{8}\overline{\mathbf{e}}_{\mathbf{v}} \\
\mathbf{e} &\longleftarrow \mathbf{e} + 2\overline{\mathbf{v}}_{\mathbf{e}} \\
\mathbf{f} &\longleftarrow \mathbf{f} + 4\overline{\mathbf{v}}_{\mathbf{f}} + 4\overline{\mathbf{e}}_{\mathbf{f}} \\
\mathbf{e} &\longleftarrow \mathbf{e} + \tfrac{1}{2}\overline{\mathbf{f}}_{\mathbf{e}} \\
\mathbf{v} &\longleftarrow \mathbf{v} + \tfrac{1}{4}\overline{\mathbf{f}}_{\mathbf{v}} + \overline{\mathbf{e}}_{\mathbf{v}}
\end{aligned}
\tag{A.3}
$$

Appendix B

Principal Component Analysis (PCA)

PCA [93] is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA can be used for dimensionality reduction in a data set by retaining those characteristics of the data set that contribute most to its variance through keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the ”most important” aspects of the data.

Fig. B.1 shows the PCA for a two-dimensional example dataset. Firstly, as shown in Fig. B.1(a), in the original coordinate system, obviously, the variation of the dataset is distributed almost evenly in both 2 axes x and y. Next, in Fig. B.1(b), PCA computes the 2 principal component c_1 and c_2, which are 2 orthogonal directions explaining the variation in the dataset in a decreasing order. By setting the mean of dataset, $\overline{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x_i}$ as the new origin and c_1 and c_2 as 2 coordinate axes, a new coordinate system can be defined as shown in Fig. B.1(c). In this new coordinate, since as much variability as possible is thereby represented in c_1, the data could be approximated by only ignoring the c_2 axis and thus



(a) (b) (c)

Fig. B.1: Principal components analysis of 2D dataset. (a) the distribution of data in the original coordinate system (b) the 2 principal component c_1 and c_2. (c) data represented in the new coordinate system.

reducing the dimensionality.

The principal components are computed from the empirical covariance matrix $\mathbf{\Sigma}$ of the dataset, which are defined as:

$$\mathbf{\Sigma} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x_i} - \overline{\mathbf{x}}) \cdot (\mathbf{x_i} - \overline{\mathbf{x}})^{\top} \tag{B.1}$$

where N is the number of samples in the dataset, and $\mathbf{x_i} - \overline{\mathbf{x}}$ is the deviation of the M-dimensional sample $\mathbf{x_i}$ from the arithmetic mean of the whole dataset $\overline{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x_i}$.

The eigen variation modes $\mathbf{u}_k$, $k = 1...min(M, N)$ are the unit eigenvectors of the covariance matrix $\mathbf{\Sigma}$ and defined by:

$$\mathbf{\Sigma}\mathbf{u}_k = \lambda_k\mathbf{u}_k \tag{B.2}$$

and

$$\mathbf{u}_k^T\mathbf{u}_k = 1 \tag{B.3}$$

where λ_k are the eigenvalues of the matrix $\mathbf{\Sigma}$ in the order so that $\lambda_k \geq \lambda_{k+1}$.

Rewriting (B.2) can get the set of linear equations has to be solved:

$$(\mathbf{\Sigma} - \lambda_k I)\mathbf{u}_k = 0 \tag{B.4}$$

where I is the identity. Note that the number of eigenvectors of a matrix is equal to its rank. Accordingly, $\mathbf{\Sigma}$ has $min(M, N)$ eigenvectors. The amount of variance described by an eigenvector $\mathbf{u}_k$ is proportional to the corresponding eigenvalue λ_k. So the eigenvector corresponding to eigenvalues with the largest absolute value describes the most significant modes of variability.

Because most of the variation can usually be explained by a relatively small number of eigenmodes t. Considering a smaller number $t < min(M, N)$ of eigenmodes, they will describe a proportion of the total variance

$$\lambda_t = \sum_{k=1}^{t} \lambda_k \tag{B.5}$$

The number t of the selected eigenmodes is chosen considering the overall proportion of variance explained in the selected t eigenmodes or by selecting eigenmodes with eigenvalues above a given minimum. After choosing t, any new object can be approximated by a weighted sum of the first t eigenmodes and the mean $\overline{\mathbf{x}}$:

$$\mathbf{x} = \overline{\mathbf{x}} + \mathbf{U}_t \mathbf{b}_t \tag{B.6}$$

where $\mathbf{b}_t = (b_1, b_2, ..., b_{t-1}, b_t,)^\top$ is the weight vector, and $\mathbf{U}_t = (\mathbf{u}_1, ..., \mathbf{u}_t)$ is the eigenvector matrix.