

1-1-2011

Multi-scale and multi-spectral shape analysis: from 2d to 3d

Zhaoqiang Lai
Wayne State University,

Follow this and additional works at: http://digitalcommons.wayne.edu/oa_dissertations



Part of the [Computer Sciences Commons](#)

Recommended Citation

Lai, Zhaoqiang, "Multi-scale and multi-spectral shape analysis: from 2d to 3d" (2011). *Wayne State University Dissertations*. Paper 355.

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.

**MULTI-SCALE AND MULTI-SPECTRAL SHAPE ANALYSIS:
FROM 2D TO 3D**

by

ZHAOQIANG LAI

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2011

MAJOR: COMPUTER SCIENCE

Approved by:

Advisor

Date

© COPYRIGHT BY
ZHAOQIANG LAI
2011
All Rights Reserved

DEDICATION

To the memories of my grandfather.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Jing Hua, for his continuous support in the Ph.D. program. Dr. Jing Hua is an excellent role model for any young researcher to emulate. During my Ph.D. study in Wayne State University these years, he has tirelessly spent numerous hours with me discussing research, teaching me to write papers, and answering my questions. He has been available to me all the time whenever I needed his feedback. This dissertation would not have been possible without him.

I am very grateful to Dr. Farshad Fotouhi, Dr. Ming Dong, and Dr. Ji Zhu for serving on my dissertation committee and giving constructive suggestions and comments on the dissertation.

Also, I thank all my colleagues in Graphics & Imaging Lab, Dr. Guangyu Zou, Dr. Yunhao Tan, Dr. Cui Lin, Chang Liu, Jiaxi Hu, Darshan Pai, and Vahid Taimouri, for having insightful discussions with me.

Last but not least, I deeply acknowledge the love and encouragement from my wife and my parents. I hope I have made them all proud.

TABLE OF CONTENTS

Dedication	ii
Acknowledgments	iii
List of Tables	vii
List of Figures	viii
Chapter 1 INTRODUCTION	1
1.1 Motivation	1
1.2 Contribution	5
1.3 Dissertation Organization	7
Chapter 2 BACKGROUND	9
2.1 Conformal Mapping	9
2.1.1 Harmonic maps	10
2.1.2 Conformal Maps	10
2.1.3 Conformal Maps by Energy Minimization	13
2.1.4 Conformal Maps by Least Squares Conformal Maps	14
2.2 Scale Invariant Feature Transform (SIFT)	14
2.2.1 Review of Related Image Methods	14
2.2.2 SIFT Matching	15
Chapter 3 SHAPE VECTOR IMAGE DIFFUSION (SVID) FRAMEWORK	18
3.1 Conformal Shape Vector Image Construction	18
3.2 Shape Vector Image Diffusion	21
3.2.1 Shape Vector Image Diffusion and Diffusion Space	21
3.2.2 Properties of Diffusion Space	24
3.3 Keypoint-based Shape Descriptors	27
3.3.1 Extrema Detection	27
3.3.2 Descriptor Construction	28

3.4	Shape Matching and Registration	30
3.5	Experiments and Applications on SVID	31
3.5.1	Repeatability Under Noise	31
3.5.2	Surface Matching	33
3.5.3	Multimodality Analysis	36
3.6	Summary	38
Chapter 4 LAPLACIAN SHAPE SPECTRUM IMPLEMENTATION AND ITS APPLICATION		40
4.1	Introduction and Motivation	40
4.2	Laplacian Shape Spectrum and Manifold Harmonics	41
4.2.1	Laplacian Shape Spectrum	42
4.2.2	Manifold Harmonics	43
4.3	Discrete Laplace-Beltrami Operator on Mesh	44
4.4	Registration Using Laplacian Shape Spectrum	46
4.4.1	Colon Registration	46
4.4.2	Piecewise Registration	49
4.5	Experiments and Results	50
4.5.1	Data Acquisition and Pre-processing	50
4.5.2	Experimental Results	51
4.6	Summary	54
Chapter 5 SCALE SPACE CONSTRUCTION USING POINT-BASED MANIFOLD HARMONICS		55
5.1	Review of Related Work	55
5.2	Point-based Laplace-Beltrami Operator	56
5.2.1	Computing the Approximation of PB-LBO	57
5.2.2	Construction of the Matrix Form for PB-LBO	58

5.3	Point-based Manifold Harmonics Transform	59
5.4	Diffusion in Spectral Domain and Feature Detection	60
5.4.1	Construct Scale Space by Diffusion in Spectral Domain	60
5.4.2	Keypoint Detection	63
5.4.3	Curvature Function of Point Clouds	66
5.5	Feature Descriptor and Matching	67
5.5.1	Feature Descriptor and Matching	67
5.5.2	Point Cloud Matching	68
5.5.3	Experiments and Results	71
5.6	Summary	74
Chapter 6	CONCLUSION	75
6.1	Contributions	75
6.2	Future Work	76
Appendix	78
Bibliography	80
Abstract	90
Autobiographical Statement	92

LIST OF TABLES

Table 3.1:	Runtime of the shape vector image (SVI) construction, feature extraction and matching.	36
Table 4.1:	Results of comparison between shape spectrum method and deformable model.	52

LIST OF FIGURES

Figure 2.1:	Edge $\{v_0, v_1\}$ connects two faces $\Gamma_\alpha, \Gamma_\beta$, and two corners α, β are against it. The edge weight $k_{\{v_0, v_1\}}$ is defined as the summation of the cotangents of these angles.	11
Figure 3.1:	Shape Vector Image. (a) shows the Igea (5002 vertices) surface and mesh; (b) and (c) show the mean curvature channel and conformal factor channel of the shape vector image representation of the Igea model; (d) is the composite shape vector image including both channels.	19
Figure 3.2:	The diffused shape vector images, consisting both curvature and conformal factor channels, of the Igea model at different diffusion scales, t , computed by the geodesic distance-weighted diffusion.	25
Figure 3.3:	Example of the 1D shape “image”.	26
Figure 3.4:	Keypoint detection in the diffusion space. (a) The Igea model with all the detected keypoints at different scales indicated by the points of different colors and sizes. (b) All the detected keypoints shown on the curvature channel of the shape vector image. (c-e), (g-i) and (k-m) show the intermediate curvature channel images of the DoDs across scales t and the detected extrema (shown by points) on the corresponding DoDs at different scales. (f), (j) and (n) show again the extrema detected at (e), (i) and (m), respectively, with the different sizes of circles indicating the sizes of scales at which these extrema are detected.	29
Figure 3.5:	A keypoint descriptor is generated by computing the gradient magnitude and orientation at every pixel around the keypoint (16×16 sample pixels). These samples are then accumulated into orientation histograms summarizing the contents over 4×4 subregions, indicated with thicker framed boxes. The right panel shows one subregion with the length of each arrow corresponding to the sum of the gradient magnitudes along that direction within the region.	30
Figure 3.6:	Repeatability of keypoint features when the Igea model is under different Gaussian noise levels. The left panel shows the Igea models (with the computed curvature colormaps) with 4% and 10% additive Gaussian noise and their corresponding shape vector images. The detected keypoints are shown in the shape vector images. The right panel shows the repeatability of the feature points extracted by our geodesic distance-weighted shape vector image diffusion method. The comparison to the conventional anisotropic and isotropic diffusion methods is demonstrated.	32

Figure 3.7:	Matching of face models with different expressions from the same subject. The left panel shows all the matched keypoints between the two surfaces. The right panel shows the scales of the matched keypoints.	33
Figure 3.8:	Matching of two different subjects' brain surfaces. (a) 10% of matched points are shown using the linked lines. (b) The overlap test on one registered brain region.	34
Figure 3.9:	The experimental results on matching arbitrary two different brain surfaces randomly selected from 20 subjects. The comparison shows that our method constantly outperforms the regular anisotropic diffusion method and the SIFT method.	35
Figure 3.10:	The multimodality image analysis pipeline. The referenced brain is used as the template SVI (TSVI), and then all other brain SVIs are registered based on this TSVI. Based on the registered shape vector images, multimodality data such as the PET and DTI, can be integrated over the SVI images to perform the multimodality analysis.	37
Figure 3.11:	Population-based PET and DTI image analysis. (a) The regions within the black contours are the detected abnormal regions in the PET texture image; (b) The regions in black are the corresponding abnormalities on the individual's brain cortical surface. (c) shows a 3D rendering of a normal DTI fiber connectivity. (d) shows the abnormal DTI map where the abnormality is contoured in red.	38
Figure 4.1:	Illustration of eigenvectors with color map on the colon surface. (a) is the first eigenvector on the colon surface; (b) is the second eigenvector on the colon surface; (c) is the third eigenvector on the colon surface; (d) is the fourth eigenvector on the colon surface.	46
Figure 4.2:	Illustration of Fiedler vector embedding and landmarks detection. The Fiedler vector gives a natural ordering of the vertex of the colon meshes in (a) and (b). The vector value has been normalized and color map. (c) and (d) show the reliable automatic detected landmarks which are rendering in yellow balls.	47
Figure 4.3:	Illustration of false positive reduction. In (a), in supine position it has a false positive which the green arrow pointing at. In (b), in prone position it confirms that it is a pseudo-polyp but not a true polyp. (c) and (d) is the virtual colonoscopy of (a) and (b). (e) and (f) show another example which has false positives in prone position and can be confirmed as pseudo-polyps in supine position.	53

Figure 5.1:	Diffusion on the camel point cloud model. From left to right are the diffused curvature function with t increasing. As the images show, as the t increases, the curvature function on the camel model becomes smoother and smoother.	63
Figure 5.2:	Illustration of <i>GPS</i> embedding. The left image shows the camel model and a red point – a reference point based on which we can calculate the analogical geometric distance from all other points. The right image shows the analogical geometric distance to the reference point with color map from red to blue. The bigger value, towards blue in the color map, means the closer distance between the current point and the reference point.	65
Figure 5.3:	Feature point detection directly on point clouds. (a) A point cloud model; (b) The detected feature points (in red) on the point cloud model; (c) The point cloud model with a different pose; (d) The feature points (in red) directly detected from point clouds.	66
Figure 5.4:	Keypoint detection on the camel model. From left to right are the keypoints detected on the model with scale t increasing. From the figure, it is clear that at the finer scales, more sophisticated feature points will be detected and in the coarse scales, the number of keypoints decreases.	66
Figure 5.5:	Shell-sector model: the numbers are the bin-indices.	69
Figure 5.6:	Illustration of the matching result. Matched points are connected by linked lines.	72
Figure 5.7:	Face matching. (a) Matching between the same subject with different expressions. (b) Matching between two different subjects. (c) and (d) are the point cloud views of (a) and (b) respectively.	72
Figure 5.8:	Statistical results. The red curve shows the inter-subject matching and the blue one shows the intra-subject matching.	73

CHAPTER 1

INTRODUCTION

Shape analysis is a fundamental aspect of many problems in computer graphics and computer vision, including surface matching and surface registration. There has been a lot of research on 3D shape analysis. The key question in shape analysis has been the choice of the shape representation scheme. Actually, in recent years a large amount of research on how to find the representation for the 3D shape has been done. Different approaches including curvature-based representations [68], regional point representations [23][54][62][12], spherical harmonic representations [26][16][17], shape distributions [49] have been proposed. However, many of these representations are not reliable and cannot perform well under such circumstance as noise, occlusion and clutter.

1.1 Motivation

Researchers and professionals dedicated to 3D field are usually looking at the 2D field to find a similar strategy for analysis purpose. In the 2D domain, extracting features from the image is one of the most important tasks in the processing of image analysis. Line segments [18], groupings of edges [36][45], and regions [5] have been proposed as features. Nevertheless, these features are not reliable and they just work well under certain circumstance. There has been recent work on developing much reliable features. One of them is to use a corner detector. Zhang et al. [77] used the Harris corner detector to identify feature locations. Harris corner detector is also used by Schmid & Mohr [58] to identify interest points for the object recognition problem. Other approaches to appearance-based recognition including eigenspace analysis [43], color histograms [63], and receptive field histograms [56] have been proposed. However, it is difficult to extend them to cluttered and partially occluded images because of their more global features. Ohba & Ikeuchi [48] successfully apply the eigenspace approach to

cluttered images by using many small local eigen-windows, but this is time consuming. Most recently, SIFT [37], scale-invariant features, has been proposed for 2D regular images and proven to be the most robust among many other local invariant feature descriptors with respect to different geometrical changes [40]. As [2] says, “SIFT [37] has been proven to be the most robust among the other local invariant feature descriptors with respect to different geometrical changes [40]”. So SIFT should be a good approach for image analysis.

It inspires us to link the 3D analysis problem to the scale space through the 2D image analysis. Here the link is the geometric maps. Geometric maps can map the 3D surface to the 2D image. This simplifies the shape-analysis problem to a 2D image-analysis problem. When constructing shape images, geometric maps, which encode the shape information of the surface patch into the 2D image, provide a solution to the mapping problem between a 3D surface and a 2D domain. In the conformal geometry theory, each 3D shape can be mapped to a 2D domain through a global optimization and the resulting map is a diffeomorphism [15][59][60][30]. Consequently the 3D shape-analysis problem can be simplified to a 2D image-analysis problem in virtue of the conformal geometric maps. These maps are stable, insensitive to resolution changes and robust to occlusion and noise. The analysis of 2D image which integrates geometric and appearance information is a better understood problem [37][3][42]. Therefore, highly accurate and efficient 3D shape analysis algorithms can be achieved by 2D image analysis by creating the shape image.

However, since SIFT was mainly developed for gray-scale images, the usage of the scale-space processing on the special shape images or shape vector images are under-explored and its performance remains unknown. As we known, SIFT is based on the scale space theory. Scale-space processing usually convolves an image with Gaussian filters, generating a sequence of images, and then the difference of successive Gaussian-blurred images are calculated to create the Difference of Gaussians (DoG) for further analysis. In order to extract the most robust and salient features to abstract the shape vector image like SIFT, we propose to create a multiscale

vector-valued diffusion space through our novel geodesic distance-weighted shape vector image diffusion. As a result, analysis of the shape vector image in its diffusion space is similar to the direct diffusion analysis of the 3D model. A valuable point here is that our computation is executed in a regular 2D domain, which is much simpler than in the 3D domain.

In the diffusion space, we can then extract distinctive features used for matching and analysis. A rich set of scale-aware features can be extracted from the diffusion space representation. Similar to the feature extraction technique in SIFT [37], our approach detects the extrema across the scales as keypoints. We then calculate the orientation histograms around the keypoints as feature descriptors, which provide distinctive bases for representing the 3D geometry of the original shape. These scale-aware geometric features can directly be used for robust matching and registration against the noises and distortions. Therefore, statistical analysis and visualization of surface properties across subjects become readily available. This is important for many real-world applications. For example, it is very useful for processing inter-subject brain surfaces from medical scans of different subjects since these surfaces exhibit the inherited physiological variances among subjects. We have conducted extensive experiments on scanned real-world surface models and real 3D human neocortical surfaces, through which we demonstrate the excellent performance of our approach in surface matching and registration, statistical analysis, and integrated visualization of the multimodality volumetric data over the shape vector image.

Although we can simplify a 3D problem to 2D problem, however, the major disadvantage of the shape vector image diffusion method is that it introduces inevitable, large distortions when mapping large, complex and topologically complicated surfaces to a canonical domain. This unwanted distortion changes the characteristics of actual 3D shapes and affects the performance in shape matching or other shape analysis. In addition, it would be very difficult to compute parametrization of certain high genus-surfaces. Here we are trying to find a diffusion method which can be applied directly to the shape without the mapping.

To find the solution for shape diffusion, we have to look back to the root of 2D SIFT. Since the essential of SIFT is the diffusion equation which can create a multi-scale space, scientists are dedicating to figuring out the solution of diffusion equation for 3D shape. In [9], the spherical harmonics is used to derive the diffusion equation for a sphere from the spectral point of view. However, that only works for the sphere because it uses the spherical harmonics. It would be nice if we have the harmonics for a arbitrary regular shape instead of the spherical harmonics only for sphere. To find the harmonics, we have to turn to the shape spectrum field which provides another shape analysis tool from the spectral point of view.

Shape spectrum is a rather new research topic in computational geometry and shape analysis [25, 29, 52, 55]. Reuter [52] defined shape spectrum as the family of eigenvalues of the Laplace-Beltrami operator on a manifold, and used it as a global shape descriptor. Lévy [29] pointed out that Laplace-Beltrami eigenfunctions are tools to understand the geometry of shapes and discussed the properties of those eigenfunctions of the Laplace-Beltrami operator. Rustamov [55] proposed a Laplace-Beltrami shape distribution based on global point signatures using the eigenfunctions and eigenvalues. Karni and Gotsman [25] used the projections of geometry on the eigenfunctions for mesh compression. The Laplace shape spectrum, which has been proved to have many good invariant properties [52], is showing tremendous power and potential in shape analysis. While we are investigating the shape spectrum, we find that the Fiedler vector from shape spectrum naturally following the shape makes it as a good tool for surface registration. As a medical application in our test case, we employed the shape spectrum to extract the shape characteristics as the signature to find the correspondent regions between the prone and supine lumen surface. The method is simple yet efficient and accurate.

In the shape spectrum filed, most recently, Vallet and Lévy [65] proposed the manifold harmonics transform on the mesh which can be considered as a filter in the spectrum domain. This type of work inspired us to combine spectrum and diffusion together to seek a new creation of scale space for shapes from the spectrum domain by integrating the manifold harmonics

with the diffusion equation which is the essential mathematical foundation for the scale space. So we propose a novel method to create a scale space for the shape by finding the solution of the diffusion equation using the manifold harmonics from the spectral point of view. Not confined on the mesh, which is reconstructed from the point cloud, in this dissertation, we use the point-based manifold harmonics to deal with the point cloud data. We rigorously derive our solution from the diffusion equation on the manifold using the point-based manifold harmonics. Utilizing the manifold harmonics, the creation of scale space on the point clouds can be achieved by the diffusion of the curvature function. Based on the multiscale structure, we can robustly detect the distinctive features by comparing the neighbors directly on the point clouds. Shape descriptors associated with the local scale information can then be defined. With these feature descriptors, the point cloud matching can be achieved through the feature descriptors matching.

1.2 Contribution

We have systematically developed the proposed framework based on the ideas from the motivations. Based on this, my research work has been fruitful in my Ph.D. study with three important contributions are presented as following that fall into the filed of shape matching and registration in both academia and medical domain:

- **Shape Vector Image Diffusion Framework (*Chapter 3*).** We present a novel and efficient surface matching and visualization framework based on the geodesic distance-weighted shape vector image diffusion. Firstly, our framework conformally maps a to-be-analyzed surface to a canonical 2D domain. The surface curvatures and conformal factors are then interpolated and encoded into the rectangular 2D domain, which we call shape vector image. Based on the shape vector image, we propose to create a multiscale vector-valued diffusion space through our novel geodesic distance-weighted shape vector image diffusion. By detecting the cross-scale extrema in the diffusion space, we can create a set of distinctive scale-aware geometric features. These robust features are well

suitable for surface matching and registration. We have conducted extensive experiments on scanned real-world surface models and real 3D human neocortical surfaces, through which we demonstrate the excellent performance of our approach in surface matching and registration, statistical analysis, and integrated visualization of the multimodality data over the shape vector image.

- **Implementation and Application of Laplacian Shape Spectrum (*Chapter 4*).** A novel surface registration method based on shape spectrum is invented and applied to colon surface registration. First of all, we show the implementation of the Laplace-Beltrami operator on the mesh. By constructing the matrix form for the Laplace-Beltrami operator, we can solve the eigen problem to get the Laplacian Shape Spectrum. The second eigenvector of the Laplacian is called the Fiedler vector and has interesting properties, and we employ the Fiedler vector to extract the shape characteristics as the signature for registration. To do the registration efficiently, we first detect reliable anatomic landmarks. With the landmarks, we register the colon surface using piecewise registration using the Fiedler vector value. We apply the algorithm to the real computed tomography colonography (CTC) datasets and the experiments demonstrate the excellent accuracy of our registration results. Furthermore, we integrate the registration component into our virtual colonoscopy software for false positive (FP) reduction, which shows that it is an excellent tool aiding polyp diagnosis in CTC.
- **Scale Space Construction Using Point-based Manifold Harmonics (*Chapter 5*).** We present a novel method to create a scale space for the point cloud from the spectral point of view by using the manifold harmonics. First of all, we show how to implement the Laplace-Beltrami operator for the point cloud data. Based on the Laplace-Beltrami operator, we have the point-based manifold harmonics. Using the point-based manifold harmonics, we rigorously derive our solution from the diffusion equation which is the essential of the scale space processing on the manifold. Built upon the point-based man-

ifold harmonics transform, we generalize the diffusion function directly on the point clouds to create the scale space. Based on the multi-scale structure, we try to detect the features as the local extrema. Since there is no connectivity on the point cloud data, we have to define a local neighborhood for the feature detection. By using the global point signature, a neighborhood can be defined around the detected features on the point cloud data. Subsequently, we can construct a feature descriptor for each feature point which is effective for matching point clouds. With these feature descriptors, the point cloud matching can be achieved through the feature descriptors matching. The performance of our framework was thoroughly evaluated through our experiments which demonstrated that our method can achieve excellent matching performance for point cloud in real-world applications.

1.3 Dissertation Organization

The dissertation is organized in the following way:

- Chapter 2: Gives definitions and detailed description of conformal mapping techniques and scale invariant feature transform. This provides a mathematical foundation on which the remainder of the thesis will be based.
- Chapter 3: Describes how shape vector image diffusion can be used for multiscale diffusion space construction. This chapter presents the detection, descriptor and matching of keypoints. Applications are given as well as experiments validating the technique.
- Chapter 4: Shows the implementation of the Laplacian shape spectrum on mesh and presents the surface registration application using the Laplacian shape spectrum.
- Chapter 5: Present a novel method to create a scale space for the point cloud from the spectral point of view by using the point-based manifold harmonics.

- Chapter 6: Offers a summary of the dissertation and points towards the future work along this research direction.

CHAPTER 2

BACKGROUND

This chapter provides a mathematical background for the problem we raise in previous chapter. This foundation will facilitate the reader’s understanding of the research and contributions put forth in the following chapters. This chapter is organized as follows. *Chapter 2.1* reviews conformal mapping, including the introduction of conformal mapping and the calculation of the conformal mapping. *Chapter 2.2* reviews scale invariant feature transform.

2.1 Conformal Mapping

Geometric maps have been widely applied in the fields of computer graphics and computer vision. Zhang and Hebert used harmonic maps to construct scalar shape images to match the surfaces [80]. In [73], harmonic maps were used to track dynamic 3D surface. However, calculating harmonic maps needs to identify the surface boundary and create the boundary mapping from 3D surfaces to the 2D domain, which becomes unreliable when there are noises and occlusions in the 3D original data. Since the interior feature points are often more robust in the 3D original data, conformal maps, which do not need boundary information, can be a natural choice to overcome the difficulty. Using several feature constraints instead of the boundary condition, conformal maps have many appealing properties [72]. For example, if the parameterization is conformal, the surface can be uniquely determined by the mean curvatures with area stretching factors defined on the parametric domain. In [19], genus zero surface conformal mapping was discussed and it was adopted for the brain surface mapping [81]. The conformal parameterization can be uniquely determined by two corresponding points. Conformal parameterization depends on the geometry itself, not the triangulation of the surfaces. Hence, conformal mapping is a viable solution for 3D shape image construction. This motivated us to encode a shape vector image using conformal mapping for surface representation. In this

section, we will review the existing work of the conformal mapping.

2.1.1 Harmonic maps

Given two genus zero meshes M and N . A map $f : M \rightarrow N$, the string energy is defined as:

$$E(f) = \langle f, f \rangle = \sum_{\{v_0, v_1\}} k_{\{v_0, v_1\}} \|f(v_0) - f(v_1)\|^2, \quad (2.1)$$

where $\{v_0, v_1\}$ is an edge connecting two neighboring vertices v_0 and v_1 .

Suppose edge $\{v_0, v_1\}$ has two adjacent faces $\Gamma_\alpha, \Gamma_\beta$, with $\Gamma_\alpha = \{v_0, v_1, v_2\}$ and $\Gamma_\beta = \{v_0, v_1, v_3\}$, as shown in Fig. 2.1; $k_{\{v_0, v_1\}}$ is defined as

$$\frac{1}{2} \left(\frac{(v_0 - v_2) \cdot (v_1 - v_2)}{|(v_0 - v_2) \times (v_1 - v_2)|} + \frac{(v_0 - v_3) \cdot (v_1 - v_3)}{|(v_0 - v_3) \times (v_1 - v_3)|} \right),$$

and the string energy obtained with this $k_{\{v_0, v_1\}}$ is called the harmonic energy.

Specially, if string constants $k_{\{v_0, v_1\}} \equiv 1$, the string energy is known as the Tutte energy.

The harmonic map can be computed by minimizing the harmonic energy using the Euler-Lagrange differential equation. With the boundary condition of the mapping f is given, $f|_{\partial M} : \partial M \rightarrow \partial N$, the solution is unique. The harmonic maps are easy to compute, however, it becomes unreliable when there are occlusions in the 3D original data. As it has to meet the boundary condition, the occlusion will cause the problem [80]. Conformal mapping, which replaces the boundary condition with feature constraints, is proposed because interior feature points are often more robust to occlusion.

2.1.2 Conformal Maps

In the theorem of differential geometry, a diffeomorphism $f : M \rightarrow N$ is conformal if and only if, for any surface patch σ_m on M , the first fundamental forms of σ_m and $\sigma_n = f \circ \sigma_m$ are proportional. Mathematically, this means that

$$f \circ ds_m^2 = \lambda ds_n^2, \quad (2.2)$$

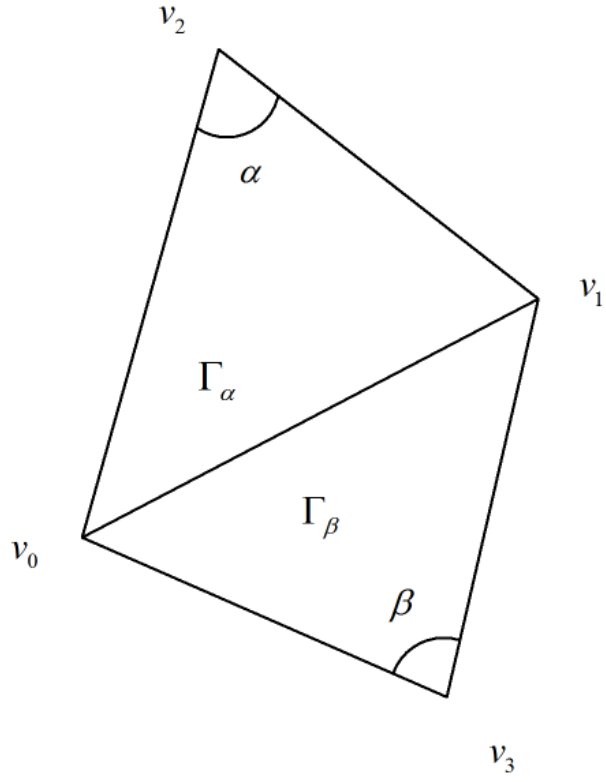


Figure 2.1: Edge $\{v_0, v_1\}$ connects two faces $\Gamma_\alpha, \Gamma_\beta$, and two corners α, β are against it. The edge weight $k_{\{v_0, v_1\}}$ is defined as the summation of the cotangents of these angles.

where λ is called the conformal factor, ds_m^2 and ds_n^2 are the first fundamental form on M and N [20].

If M and N are surfaces, a diffeomorphism $f : M \rightarrow N$ is said to be conformal if , whenever f takes two intersecting curves γ_m and $\tilde{\gamma}_m$ on M to curves γ_n and $\tilde{\gamma}_n$ on N , the angle of intersection of γ_m and $\tilde{\gamma}_m$ is equal to the angle of intersection of γ_n and $\tilde{\gamma}_n$. In short, f is conformal if it preserves angles.

It is well known that any genus zero surface can be mapped conformally onto the sphere and any local portion thereof onto a disk. This mapping, a conformal equivalence, is one-to-one, onto, and angle-preserving. Moreover, the elements of the first fundamental form remain unchanged, except for a scaling factor (the so-called Conformal Factor). For this reason, conformal mappings are often described as being similarities in the small region. Since the cortical surface of the brain is a genus zero surface, conformal mapping offers a convenient method to retain local geometric information, when mapping data between surfaces.

Another nice property of conformal mappings is their connection to complex function theory. Suppose a mapping $f : M \rightarrow N$ between two planar Riemann surfaces, which can be considered as a function of a complex variable, $n = f(m)$. A conformal map thus satisfies the Cauchy-Riemann equations, which, with $n = u + iv$ and $m = x + iy$, are

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (2.3)$$

Noticing that by differentiating one of these equations with respect to x and the other with respect to y , we obtain the two Laplace equations

$$\Delta u = 0, \quad \Delta v = 0, \quad (2.4)$$

where

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (2.5)$$

is the Laplace operator.

Any mapping which satisfies these two Laplace equations is called a harmonic mapping. Thus a conformal mapping is also harmonic. A harmonic map can either be viewed as the solution to Eq. 2.4, or as the minimizer of the energy (Eq. 2.1) over the surface.

2.1.3 Conformal Maps by Energy Minimization

Suppose K representing the simplicial complex; then we have these kinds of definitions here [19]:

Definition 1: The piecewise Laplacian is the linear operator $\Delta_{PL} : C^{PL} \rightarrow C^{PL}$ on the surface of piecewise linear functions on K , defined as

$$\Delta_{PL}(f) = \sum_{\{v_0, v_1\}} k_{\{v_0, v_1\}} (f(v_0) - f(v_1)). \quad (2.6)$$

Definition 2: For a map $\vec{f} : M \rightarrow \mathbb{R}^3$, $\vec{f} = (f_0, f_1, f_2)$, $f_i \in C^{PL}$, $i = 0, 1, 2$, the energy is defined as the norm of \vec{f}

$$E(\vec{f}) = \sum_{i=0}^2 E(f_i). \quad (2.7)$$

Definition 3: For a map $\vec{f} : M \rightarrow \mathbb{R}^3$, the piecewise Laplacian of \vec{f} is

$$\Delta_{PL} \vec{f} = (\Delta_{PL} f_0, \Delta_{PL} f_1, \Delta_{PL} f_2). \quad (2.8)$$

The minimization of the energy $E(\vec{f})$ can be solved by the steepest descent algorithm

$$\delta \vec{f} = -D \vec{f} \times \delta t, \quad (2.9)$$

where $D \vec{f}$ is the absolute derivative. The detail of the algorithm can be referred to [19].

2.1.4 Conformal Maps by Least Squares Conformal Maps

The Least Squares Conformal Map (LSCM) parameterization algorithm [30] generates a discrete approximation of a conformal map by adding more constraints. Here we give a brief description.

Given a discrete 3D surface mesh M and a smooth target mapping $f : M \rightarrow (u, v)$, then, as described in previous section, f is conformal if and only if the Cauchy-Riemann equation holds true on the whole of M . However, in general this conformal condition cannot be strictly satisfied on the whole triangulated surface M , so the conformal map is constructed in the least squares sense:

$$C(M) = \sum_{d \in M} \left\| \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right\|^2 A(d), \quad (2.10)$$

where d is a triangle on the mesh M with the area $A(d)$. The least squares minimization problem in Eq. 2.10 can be efficiently solved using the Conjugate Gradient Method.

2.2 Scale Invariant Feature Transform (SIFT)

First we will review the related 2D image feature extraction techniques. Then we detail the SIFT algorithm.

2.2.1 Review of Related Image Methods

First, let me give a short review of the related works on 2D image methods. After mapping the 3D surface to the 2D image plane, extracting features from the shape image becomes one of the most important tasks. Generally speaking, the analysis of 2D image is a better understood problem [8, 3, 42]. However, the analysis of the shape image that integrates geometric and appearance information has its own special challenges mainly due to the non-uniform sampling and different pixel properties (i.e., geometric characteristics instead of grey-scale intensities). Hence, conventional image analysis techniques may not work well. For example, line segments [18], groupings of edges [36, 45], and regions [5] have all been proposed as features for image matching. Nevertheless, these feature extraction techniques are not reliable for shape images

due to the aforementioned special characteristics of shape images and they only work well under certain circumstances. Recently, there has been a great deal of research work on developing more reliable features for conventional images. One of them is to take advantage of corner detectors. Zhang et al. [77] used the Harris corner detector to identify feature locations. Harris corner detector was also used by Schmid and Mohr [58] to identify interest points for the object recognition problem. Other approaches have been proposed for appearance-based matching, including eigenspace analysis [43], color histograms [63], and receptive field histograms [56]. However, it is difficult to extend them to match variant inter-subject images because of their more global features. Ohba and Ikeuchi [48] successfully applied the eigenspace approach to cluttered images by using many small local eigen-windows, but this is very time consuming and not practical. Most recently, SIFT, scale-invariant features, has been proposed [37] for 2D regular images and proven to be the most robust among many other local invariant feature descriptors with respect to different geometrical changes [40, 2].

2.2.2 SIFT Matching

From the fact [34] that the scale-space representation $L(x, t)$ satisfies the diffusion equation

$$\frac{\partial L(x, t)}{\partial t} = \text{div}(\nabla L(x, t)) = \nabla^2 L(x, t), \quad (2.11)$$

where the Laplacian on the right side is taken only with respect to the x variables ($x \in R^n$), it follows that the Laplacian can also be computed as the limit case of the difference between two Gaussian smoothed images. So in SIFT, the image is convolved with Gaussian filters generating a sequence of images, and then the difference of successive Gaussian-blurred images are calculated to create the Difference of Gaussians (DoG). Keypoints are then taken as maxima/minima of the DOG at each scales. Specifically, a DoG image $D(x, y, \sigma)$ is given by

$$D(x, y, \sigma) = L(x, y, k_{i+1}\sigma) - L(x, y, k_i\sigma), \quad (2.12)$$

where $L(x, y, k_i\sigma)$ and $L(x, y, k_{i+1}\sigma)$ are the original image $I(x, y)$ convolved with the Gaussian blur $G(x, y, k\sigma)$ at scale $k_i\sigma$ and $k_{i+1}\sigma$, i.e.,

$$L(x, y, k_i\sigma) = G(x, y, k_i\sigma) * I(x, y) \quad (2.13)$$

and

$$L(x, y, k_{i+1}\sigma) = G(x, y, k_{i+1}\sigma) * I(x, y), \quad (2.14)$$

where

$$G(x, y, k_i\sigma) = \frac{1}{2\pi k_i\sigma} e^{-(x^2+y^2)/2k_i\sigma} \quad (2.15)$$

and

$$G(x, y, k_{i+1}\sigma) = \frac{1}{2\pi k_{i+1}\sigma} e^{-(x^2+y^2)/2k_{i+1}\sigma}. \quad (2.16)$$

Once DoG images have been obtained, keypoints are identified as local minima/maxima of the DoG images across scales. Detection of the interest points at the extrema of a difference of Gaussian pyramid of the input image is performed. By comparing each pixel in the DoG images to its eight neighbors at the same scale and nine corresponding neighboring pixels in each of the neighboring scales, the detection can be done and all the maximum or minimum will be selected as candidate keypoints.

Next, each keypoint is assigned orientation based on local image gradient directions. Orientation assignment is the key step in achieving invariance to rotation as the keypoint descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation after the interesting points are found. For an image sample $L(x, y)$ at scale σ , the gradient magnitude, $m(x, y)$, and orientation, $\theta(x, y)$, are precomputed using pixel differences:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2.17)$$

and

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \right). \quad (2.18)$$

Using above equations, the magnitude and direction are calculated for every pixel in a neighboring region around the keypoint in the Gaussian-blurred image L . An orientation histogram with 36 bins is formed, with each bin covering 10 degrees. Once the histogram is filled, the orientations which have largest value and those which are within 80% of the largest are assigned to the keypoint.

The feature descriptor is computed as a set of orientation histograms on the (4×4) pixel neighborhoods, each of which have 8 orientation bins, around the interesting point. So it uses this $4 \times 4 \times 8 = 128$ elements vector for the feature descriptor. This vector is then normalized to unit length in order to enhance invariance to affine changes in illumination.

Then the SIFT matching is achieved by matching the feature descriptor. The matching is calculated based on Euclidean distance of the feature vectors.

CHAPTER 3

SHAPE VECTOR IMAGE DIFFUSION (SVID)

FRAMEWORK

This chapter presents a novel and efficient surface matching and visualization framework through the geodesic distance-weighted shape vector image diffusion. Based on conformal geometry, our approach can uniquely map a 3D surface to a canonical rectangular domain and encode the shape characteristics (e.g., mean curvatures and conformal factors) of the surface in the 2D domain to construct a geodesic distance-weighted shape vector image, where the distances between sampling pixels are not uniform but the actual geodesic distances on the manifold. Through the novel geodesic distance-weighted shape vector image diffusion presented in this chapter, we can create a multiscale diffusion space, in which the cross-scale extrema can be detected as the robust geometric features for the matching and registration of surfaces. Therefore, statistical analysis and visualization of surface properties across subjects become readily available. The experiments on scanned surface models show that our method is very robust for feature extraction and surface matching even under noise and resolution change. We have also applied the framework on the real 3D human neocortical surfaces, and demonstrated the excellent performance of our approach in statistical analysis and integrated visualization of the multimodality volumetric data over the shape vector image.

3.1 Conformal Shape Vector Image Construction

A good shape image should be able to fully represent the geometric characteristics of a given surface, and thus serves as a domain for indexing other heterogenous attributes. Thus, a 3D surface can be converted to a multidimensional vector image for effective processing. We employ conformal mapping to accomplish the task. In the theorem of differential geometry, a diffeomorphism $f : M \rightarrow N$ is conformal if and only if, for any surface patch σ_m on M ,

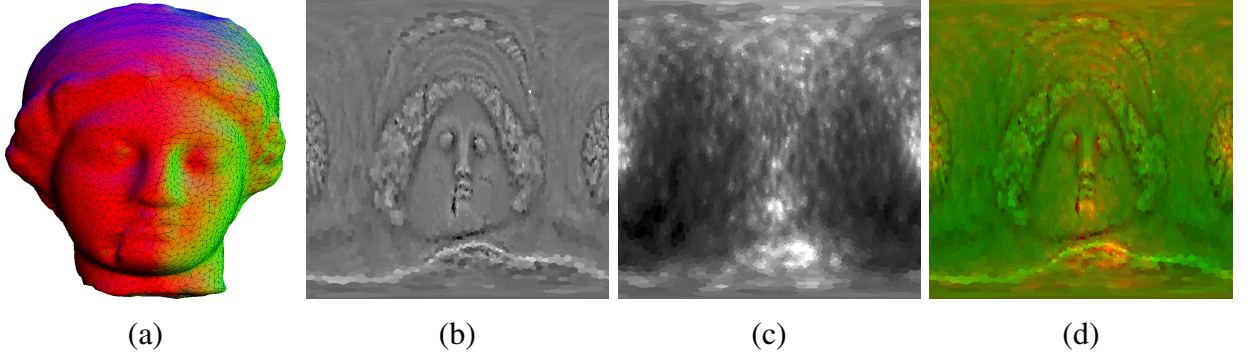


Figure 3.1: Shape Vector Image. (a) shows the Igea (5002 vertices) surface and mesh; (b) and (c) show the mean curvature channel and conformal factor channel of the shape vector image representation of the Igea model; (d) is the composite shape vector image including both channels.

the first fundamental forms of σ_m and $\sigma_n = f \circ \sigma_m$ are proportional. Mathematically, this means that $f \circ ds_m^2 = \lambda ds_n^2$, where λ is called the conformal factor, ds_m^2 and ds_n^2 are the first fundamental form on M and N . If M and N are surfaces, a diffeomorphism $f : M \rightarrow N$ is said to be conformal if, whenever f takes two intersecting curves γ_m and $\tilde{\gamma}_m$ on M to curves γ_n and $\tilde{\gamma}_n$ on N , the angle of intersection of γ_m and $\tilde{\gamma}_m$ is equal to the angle of intersection of γ_n and $\tilde{\gamma}_n$. In short, f is conformal if it preserves angles.

In order to match 3D shapes accurately and efficiently, we develop a 2D representation, shape vector image, using conformal mapping. Given a surface patch M , its conformal image I_c can be created using conformal mapping. There is one-to-one correspondence between the vertices in M and the vertices in I_c . Based on the shape attributes computed at each vertex of M , attribute values can be interpolated and computed for each pixel of the conformal shape vector image. In practice, we compute the conformal parameterization by a nonlinear optimization method carried out in the tangential space of a sphere as proposed in [19].

Suppose K denotes the simplicial complex and there is a piecewise linear embedding $\vec{l} : |K| \rightarrow \mathbf{R}^3$. Then a triangular mesh can be represented as (K, \vec{l}) . For the purpose of implementation, surfaces are usually approximated by triangular meshes. We use p, q to denote the vertices and $\{p, q\}$ to denote the edge spanned between p and q . The surface and its para-

metric domain are modeled as piecewise linear functions \vec{f} and \vec{g} in accordance with (K, \vec{l}) , respectively. The mean curvature at vertex q is estimated as in [39] by

$$H(q) = \frac{1}{4A} \left\| \sum_{p \in N_1(q)} (\cot \alpha_{p,q} + \cot \beta_{p,q}) (\vec{f}(q) - \vec{f}(p)) \right\|^2, \quad (3.1)$$

where $\alpha_{p,q}$ and $\beta_{p,q}$ are the two opposite angles of edge p, q in the two triangles sharing this edge, and $N_1(q)$ is the set of 1-ring neighbor vertices of vertex q . A is the area of the associated surface patch (so-called finite volume in Mechanics), which is given by

$$A(q) = \frac{1}{8} \sum_{p \in N_1(q)} (\cot \alpha_{p,q} + \cot \beta_{p,q}) \|\vec{f}(q) - \vec{f}(p)\|^2, \quad (3.2)$$

under the condition that the triangles in the 1-ring neighbors are non-obtuse. In case of obtuse triangles, refer to [39] for solutions. Following this path, we define the discrete conformal factor operator as

$$\lambda(q) = \frac{A_g(q)}{A_f(q)}, \quad (3.3)$$

where the $A_f(q)$ and $A_g(q)$ are the averaging areas for each homotopic vertex q on surface \vec{f} and \vec{g} , respectively.

As conformal surface representation $S(u, v)$ is parameterized by conformal parameters (u, v) on a domain D , the conformal factor function, $\lambda(u, v)$, and mean curvature function, $H(u, v)$, defined on D satisfy the Gauss and Codazzi equation. If $\lambda(u, v)$ and $H(u, v)$ are given together with the boundary conditions, $S(u, v)$ can be uniquely reconstructed. Since the mean curvature and the conformal factor are two important attributes, we assign these two attributes to $\vec{I}_c(u, v)$ to construct a vector image I , where the pixel attributes are represented by a vector $[H, \lambda]^\top$. Other features such as normal and texture can be considered if necessary. We use barycentric interpolation for sampling $\vec{I}_c(u, v)$ to I . Fig. 3.1 shows the Igea surface model with 5002 vertices (Fig. 3.1(a)) and its corresponding mean curvature channel (Fig. 3.1(b))

and conformal factor channel (Fig. 3.1(c)). The composite shape vector image is shown in Fig. 3.1(d).

3.2 Shape Vector Image Diffusion

As we known, SIFT is based on the scale space theory. Scale-space processing usually convolves an image with Gaussian filters, generating a sequence of images, and then the difference of successive Gaussian-blurred images are calculated to create the Difference of Gaussians (DoG) for further analysis. Since scale-space theory was mainly developed for gray-scale images, the usage of the scale-space processing on the special shape images or shape vector images are under-explored and its performance remains unknown.

Since a surface can be represented as a unique shape vector image composed of conformal factors and curvatures, many algorithms suitable for image computing may be used for feature extraction from this type of images. For the purpose of matching and visualization of cross-subject data, the main task is to find the stable keypoints or regions and their local image features for alignment. Since the shape vector image representation that we propose consists of the mean curvature and area distortion, it provides important signature of the local geometry, which is transformation invariant and suitable for shape matching. This section describes a novel diffusion-based algorithm to extract distinctive features from the shape vector images. Through the geodesic distance-weighted shape vector image diffusion, we can identify the robust keypoints and their scales from the computed diffusion extrema, which are suitable for the matching purpose.

3.2.1 Shape Vector Image Diffusion and Diffusion Space

Our shape vector image is a multichannel image. The simplest way to perform the diffusion filtering of the shape vector image is to deal with each channel separately and independently. However, this method leads to an undesirable effect that edges may be formed at different locations for each channel since the curvature and conformal factor channels must take effect simultaneously in order to accurately determine the local geometry. In our framework, we

employs a diffusivity g which combines information from all channels. For a vector image $I = (I_1, I_2, \dots, I_m)^\top$, the diffusion is performed by

$$\frac{\partial I_k}{\partial t} = \text{div}(g \nabla I_k) \quad (k = 1, \dots, m), \quad (3.4)$$

where div indicates the divergence operator, ∇ is the gradient operator, and $g(x) = \frac{1}{\sqrt{1+(\frac{x}{l})^2}}$ (l is a constant). For the case in which g is a constant for a specific channel I_k , it reduces to the isotropic heat diffusion equation, $\frac{\partial I_k}{\partial t} = c \Delta I_k$, where Δ is the Laplacian operators and the solution is Gaussian smoothing. However, Gaussian smoothing has a typical disadvantage, especially for the shape vector image. Gaussian smoothing does not only reduce noise, but also blurs important geometric features such as sharp edges, hence making them harder to be identified.

To solve the problem, we propose to perform geodesic distance-weighted inhomogeneous linear diffusion of the shape vector image,

$$\frac{\partial I_k}{\partial t} = \text{div}\left(g\left(\|\nabla f_{I_k}\|\right)\nabla I_k\right), \quad (3.5)$$

where I_k is the actual diffused image, f_{I_k} is the original image and g is the diffusivity function. For a specific channel $P = I_k$, the numerical solution for Eq. 3.5 can be computed, similar to the 4-nearest-neighbors discretization [51], as follows,

$$P_{i,j}^{t+1} = P_{i,j}^t + \rho[c_N \cdot \nabla_N P + c_S \cdot \nabla_S P + c_E \cdot \nabla_E P + c_W \cdot \nabla_W P], \quad (3.6)$$

where $0 \leq \rho \leq 1/4$, N, S, E, W are the subscripts for the North, South, East, and West, t is the scale, i and j are the indices of the image pixel. Since the shape vector image encodes geodesic distance information, the symbol ∇ is defined as follows with the consideration of geodesic distances:

$$\begin{aligned}
\nabla_N P_{i,j} &= \frac{P_{i-1,j} - P_{i,j}}{GeoD[N]}, & \nabla_S P_{i,j} &= \frac{P_{i+1,j} - P_{i,j}}{GeoD[S]}, \\
\nabla_E P_{i,j} &= \frac{P_{i,j+1} - P_{i,j}}{GeoD[E]}, & \nabla_W P_{i,j} &= \frac{P_{i,j-1} - P_{i,j}}{GeoD[W]},
\end{aligned} \tag{3.7}$$

where

$$GeoD[N] = GeoD([i-1, j], [i, j]), \quad GeoD[S] = GeoD([i+1, j], [i, j]),$$

$$GeoD[E] = GeoD([i, j+1], [i, j]), \quad GeoD[W] = GeoD([i, j-1], [i, j]),$$

and $GeoD([pixelA], [pixelB])$ is the normalized geodesic distance between the $pixelA$ and $pixelB$ on the manifold, which is normalized by dividing the averaged $GeoD$ over the image.

And the c is defined as:

$$c_{N_{i,j}} = g(|\nabla_N f_{I_{i,j}}|), \quad c_{S_{i,j}} = g(|\nabla_S f_{I_{i,j}}|),$$

$$c_{E_{i,j}} = g(|\nabla_E f_{I_{i,j}}|), \quad c_{W_{i,j}} = g(|\nabla_W f_{I_{i,j}}|),$$

where $\nabla_N f_{I_{i,j}}$, $\nabla_S f_{I_{i,j}}$, $\nabla_E f_{I_{i,j}}$ and $\nabla_W f_{I_{i,j}}$ are computed by Eq. 3.7. Therefore, the final numerical solution is

$$P_{i,j}^{t+1} = P_{i,j}^t + \rho \left[\frac{c_N \cdot \nabla_N P + c_S \cdot \nabla_S P}{GeoD_{NS}} + \frac{c_E \cdot \nabla_E P + c_W \cdot \nabla_W P}{GeoD_{EW}} \right], \tag{3.8}$$

where

$$GeoD_{NS} = \frac{GeoD[N] + GeoD[S]}{2},$$

$$GeoD_{EW} = \frac{GeoD[E] + GeoD[W]}{2}.$$

Using this numerical solution, we can construct a discrete geometric diffusion space which

encodes the surface geometric information.

Particularly for our shape vector image scheme, our approach sums up the diffusivity of each channel to a common diffusivity. This may be regarded as collecting the contrast information of all channels. Thus, for a shape vector image $I = (I_1, I_2, \dots, I_m)^\top$, the vector diffusion is performed by

$$\frac{\partial I_k}{\partial t} = \text{div} \left(g \left(\sum_{n=1}^m \|\nabla f_{I_n}\| \right) \nabla I_k \right) \quad (k = 1, \dots, m). \quad (3.9)$$

By solving Equations 3.9, we obtain the diffusion space,

$$\begin{pmatrix} \mathbf{I}^{t_0} & \mathbf{I}^{t_1} & \dots & \mathbf{I}^{t_n} \end{pmatrix} = \begin{pmatrix} I_1^{t_0} & I_1^{t_1} & \dots & I_1^{t_n} \\ I_2^{t_0} & I_2^{t_1} & \dots & I_2^{t_n} \\ \vdots & \vdots & \ddots & \vdots \\ I_m^{t_0} & I_m^{t_1} & \dots & I_m^{t_n} \end{pmatrix}, \quad (3.10)$$

which is a sequence of shape vector images with t as the scale in a matrix format, i.e., each row of the matrix is the sequence images of a specific channel with t as the scale, and each column of the matrix is the vector image at a specific scale t . Fig. 3.2 shows the diffused shape vector images of the Igea model with increasing diffusion scale t .

3.2.2 Properties of Diffusion Space

Suppose a mapping $f : M \rightarrow I$ and the diffusion $D : I \rightarrow S$, where M is a 3D surface, I is the shape vector image and S is the diffusion space. In this section, we firstly show that D satisfies the criteria of the multiple scale descriptions. Secondly, we show that $f \circ D$, together, creates the diffusion space appropriate for scale-space processing of the 3D surface geometry.

The diffusion space construction is to create a multiscale “semantically meaningful” descriptions of images. As we know, a scale-space representation must have the property that no spurious detail will be generated passing from finer to coarser scales. This is so-called causal-

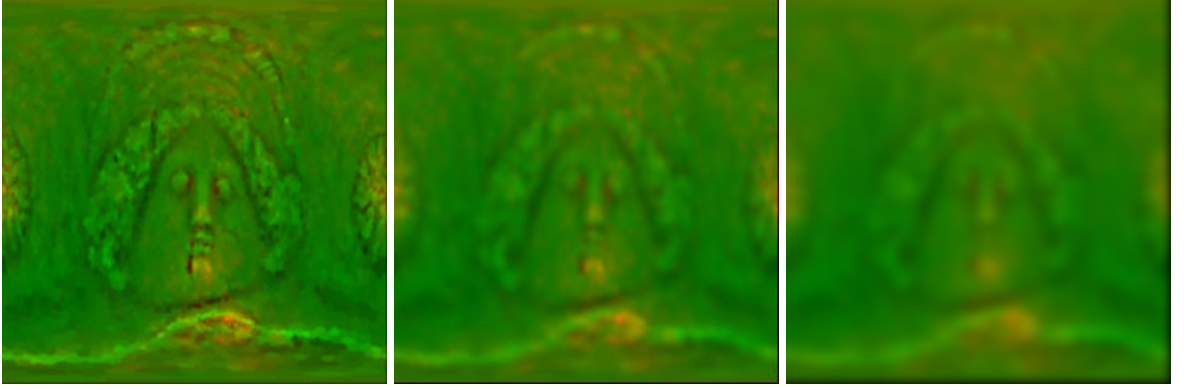


Figure 3.2: The diffused shape vector images, consisting both curvature and conformal factor channels, of the Igea model at different diffusion scales, t , computed by the geodesic distance-weighted diffusion.

ity, which means any feature at a coarse level of resolution is required to possess a “cause” at a finer level of resolution although the reverse need not be true. The causality criterion can be established by showing the used diffusion equation satisfies the maximum principle, that is to say, all the maxima of the solution of the equation in space and time belong to the initial condition (i.e., the original image). A proof of the maximum principle for the diffusion equation can be found in [51, 46]. Therefore, for the diffusion D , the satisfaction of the maximum principle leads to the satisfaction of the causality for the diffusion space. Consequently, D satisfies the criteria of the multiple scale descriptions.

For our shape vector image, we use the geodesic distance-weighted method, in which the distance can be retrieved by the mapping, f , for the computation of the diffusion to construct a geometric diffusion space. Therefore, constructing and analyzing the geometric diffusion space is similar to analyzing a direct diffusion space of the 3D surface. As a result, $f \circ D$ is able to construct a multiple scale space and multiscale descriptions for the 3D surface.

The geodesic distance-weighted anisotropic diffusion has the advantages of preserving and identifying true features as well as preventing dislocated false features in the diffusion space when taking the actual geodesic distance as a-priori information. Fig. 3.3 shows an illustrative one-dimensional example, where the curve is the actual shape object and the line segment is the

1D shape “image”. The sampling vertices (from v_1 to v_6) on the curve are mapped to the pixels (from p_1 to p_6) on the 1D shape “image” shown at the bottom. As we can see, the sampling vertices of the curve, v_2 and v_3 , have relatively high curvature values while other sampling vertices have low curvatures.

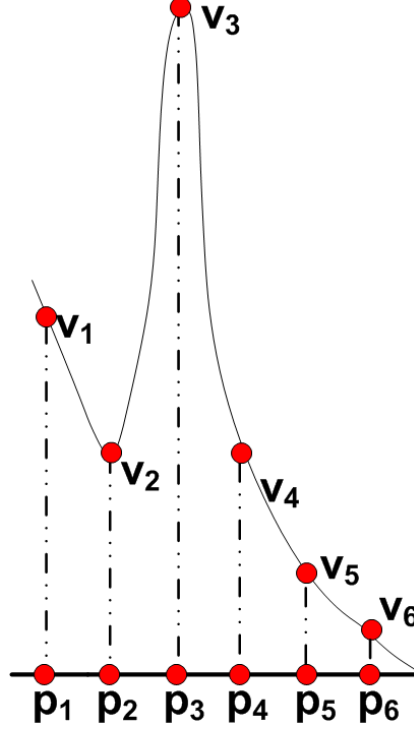


Figure 3.3: Example of the 1D shape “image”.

Without considering the geodesic distance among the pixels, p_4 's curvature value will be increased while moving from finer to coarser scales. For example, at the immediately next level, p_3 will boost the diffusion of p_4 since its distance to p_4 is considered the same as the distance between p_4 and p_5 (a unit length). The situation will get worse when diffusing further since p_2 will also dramatically affect the diffusion of p_4 . However, v_2 is far away from the diffusion vertex, v_4 . In fact, p_2 is not supposed to have significant influence on p_4 . On the other side, it is also difficult to preserve the high curvature values for p_2 and p_3 . Therefore, the procedure will dislocate keypoints when moving from finer to coarser scales. So the keypoints detected at a coarse scale do not give the correct location in the original image, which will

result in instability and incorrect matching. The situation is much improved when considering geodesic distances in the diffusion.

3.3 Keypoint-based Shape Descriptors

3.3.1 Extrema Detection

The maximum principle states that all the maxima of the solution of the equation in space and time belong to the initial condition (i.e., the original image). Therefore, we propose to detect the extrema across the diffusion space as our keypoints since they are most robust points at the specific scales which are able to represent collectively the original image. We sample the diffusion space by computing a sequence of diffused shape vector images at discrete scales, t . For each diffusion scale, we use Eq. 3.9 to calculate its diffused images which can be expressed in a matrix form like Eq. 3.10.

In order to extract the cross-scale extrema, we compute the Difference of Diffusion (DoD) using the following vector-based equation,

$$\mathbf{DoD}^{t_i} = \mathbf{I}^{t_{i+1}} - \mathbf{I}^{t_i} \quad (i = 0, \dots, n-1). \quad (3.11)$$

Consequently, we can obtain,

$$\begin{pmatrix} \mathbf{DoD}_1 \\ \mathbf{DoD}_2 \\ \vdots \\ \mathbf{DoD}_m \end{pmatrix} = \begin{pmatrix} DoD_1^{t_0} & DoD_1^{t_1} & \dots & DoD_1^{t_{n-1}} \\ DoD_2^{t_0} & DoD_2^{t_1} & \dots & DoD_2^{t_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ DoD_m^{t_0} & DoD_m^{t_1} & \dots & DoD_m^{t_{n-1}} \end{pmatrix}. \quad (3.12)$$

Fig. 3.4(c-e), (g-i), and (k-m) show the computed intermediate curvature channel images of the DoDs across scales.

Once DoD vector images have been obtained, keypoints are identified as local minima/maxima of the DoD images across scales t . For each channel \mathbf{DoD}_i , it is done by comparing each pixel

in the $\mathbf{DoD}_i^{t_j}$ images to its eight neighbors at the same scale t_j and nine corresponding neighboring pixels in each of the neighboring scales t_{j-1} and t_{j+1} . If the pixel value is the maximum or minimum among all compared pixels, it is selected as a keypoint. This algorithm is carried out through all the channels of the vector images: \mathbf{DoD}_i , ($i = 1, \dots, m$). The maxima and minima found in all the channels will be considered as the keypoints. Fig. 3.4(a) and (b) show all the detected keypoints on the Igea model. Fig. 3.4(c-e), (g-i) and (k-m) show the detected extrema (shown with points) on the corresponding DoDs at different scales. Fig. 3.4(f), (j) and (n) show the scale sizes, at which the extrema in (e), (i) and (m), respectively, are detected, with the corresponding sizes of circles. One valuable point is that the detected keypoints have the associated scales computed by the algorithm, which are very important to construct scale-aware feature descriptors.

3.3.2 Descriptor Construction

After localizing the keypoints, feature descriptors are built to characterize these points at the scales where they are identified. These descriptors should contain the necessary distinct information for their corresponding keypoints. In our framework, the descriptor is calculated channel by channel.

For each channel, the local gradient-orientation histograms of the same-scale neighboring pixels of a keypoint are used as the key entries of the descriptor. In this work, we construct a keypoint descriptor with 4×4 subdescriptors computed from a 16×16 sample pixel array, which is shown on the left side of Fig. 3.5. That is to say, a feature descriptor is computed as a set of orientation histograms on 4×4 pixel neighborhoods or subregions. The coordinates of the subdescriptors and the gradient orientations are rotated relative to the keypoint orientation (defined by the gradient vector at the keypoint location) so that it can achieve orientation invariance. One of the subdescriptors is shown on the right panel of Fig. 3.5, which gives eight directions of the orientation histogram with the length of each arrow corresponding to the magnitude of that histogram entry. Since the descriptor is computed with a 4×4 array of

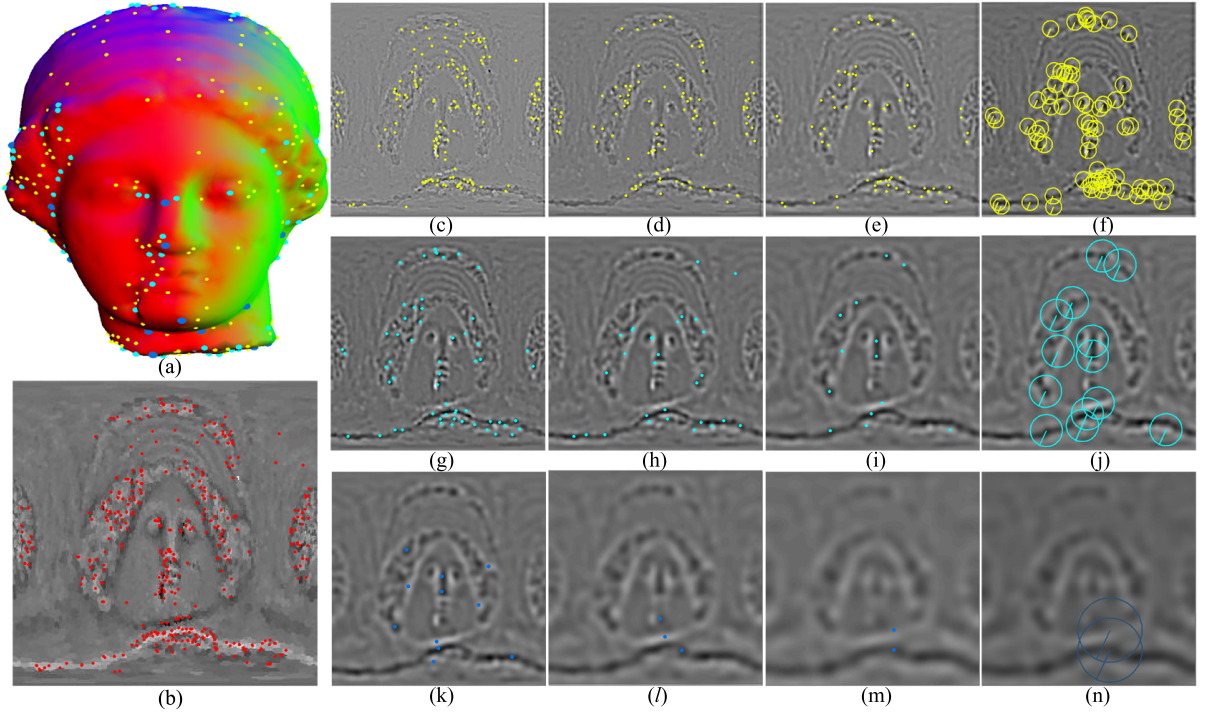


Figure 3.4: Keypoint detection in the diffusion space. (a) The Igea model with all the detected keypoints at different scales indicated by the points of different colors and sizes. (b) All the detected keypoints shown on the curvature channel of the shape vector image. (c-e), (g-i) and (k-m) show the intermediate curvature channel images of the DoDs across scales t and the detected extrema (shown by points) on the corresponding DoDs at different scales. (f), (j) and (n) show again the extrema detected at (e), (i) and (m), respectively, with the different sizes of circles indicating the sizes of scales at which these extrema are detected.

histograms with 8 orientation bins in each, this leads to a feature vector with $4 \times 4 \times 8 = 128$ elements.

In the case of an m channel vector image, a keypoint has m descriptors which are combined as a vector, $\mathbf{des} = [\mathbf{des}_1, \mathbf{des}_2, \dots, \mathbf{des}_m]^\top$, where m is the dimension of the vector image. Hence, the descriptor \mathbf{des} of a keypoint in the shape vector image is a $m \times 128$ dimension-based vector. This descriptor will be used for matching, and all the descriptors computed for all the keypoints form a feature descriptor database, which abstracts the original surface with a small number of robust keypoints and their local descriptors. The robust keypoints and constructed local shape descriptors together are well suitable for the matching purpose as demonstrated by our experiments in *Chapter 3.5*.

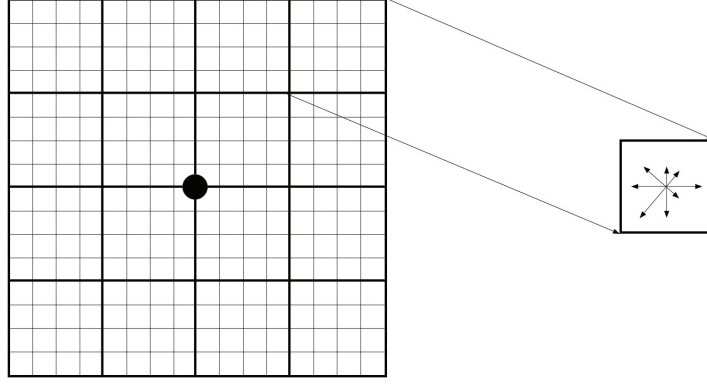


Figure 3.5: A keypoint descriptor is generated by computing the gradient magnitude and orientation at every pixel around the keypoint (16×16 sample pixels). These samples are then accumulated into orientation histograms summarizing the contents over 4×4 subregions, indicated with thicker framed boxes. The right panel shows one subregion with the length of each arrow corresponding to the sum of the gradient magnitudes along that direction within the region.

3.4 Shape Matching and Registration

In our framework, shape matching is to match the keypoints and their associated scale-aware local shape descriptors among different objects. Since the keypoints detected from the diffusion space are very reliable feature points presented in the original surface, matching these keypoints with thin-plate spline deformation will lead to accurate registration of the entire surfaces as well.

Descriptor matching is performed for a keypoint by comparing the distance from its constructed local descriptor to the descriptor of its closest matched point (DIS_{CN}) with the distance from the keypoint descriptor to the descriptor of its second-closest matched point (DIS_{SCN}) found on the to-be-matched object. The distance of two descriptors, **des1** and **des2** which are m dimension vectors, is calculated by, $DIS = \sum_{i=1}^m \| \mathbf{des1}_i - \mathbf{des2}_i \|$. Once the DIS_{CN} and the DIS_{SCN} are found, the DIS_{CN} and the DIS_{SCN} are compared to decide whether they are matched or not. The judge function for the comparison is

$$threshold \times DIS_{CN} \leq DIS_{SCN}. \quad (3.13)$$

If this inequality holds, the points are matched; Otherwise, they are not matched. This inequality ensures that only distinctive keypoints having prominent similarity are matched. Since the 3D data can be coarsely aligned easily through affine transformations during the preprocessing, we can use the uniform subdivision grid to speed up the matching. The Euclidean distance bound (ED) of two potentially matched keypoints is calculated and can be used in efficiently finding the closest and the second-closest matched points within grids.

After finding all the matched points, registration can be achieved using thin-plate splines deformation with the matched points as point constraints [83]. In our shape vector image registration, the keypoints (x_i, y_i) are taken as landmarks and $V = (x'_i, y'_i)|_i$ is a set of the matched keypoints on the other shape vector image. After computing the thin-plate splines with the above point constraints, the deformation function $f(x, y) = [f_x(x, y), f_y(x, y)]$ can be obtained to map each point (x_i, y_i) to its homolog (x'_i, y'_i) . The other unconstrained areas will follow the deformation. At the end, we can register the two shape vector images (i.e., the two 3D surfaces).

3.5 Experiments and Applications on SVID

To evaluate the proposed approach, we have conducted extensive experiments. We have applied our algorithm on real scanned face models and human neocortex surfaces extracted from high-resolution MR scans. The surface matching is demonstrated first, followed by the application of the framework in the multimodality image analysis and visualization. Our system is implemented with C++ for the computationally intensive algorithms and VTK/OpenGL for rendering and visualization. The experiments are conducted on a Dell Precision Workstation T7400, which has a Xeon CPU with Quad Cores and 4GB RAM.

3.5.1 Repeatability Under Noise

We have tested the repeatability of keypoints detection with noises. The Igea surface model is added up to 10% Gaussian noise directly on the mesh. The perturbed surfaces with different noise levels are converted to the shape vector images and then the keypoints are detected with

our geodesic distance-weighted diffusion method. We compare the repeatability of the detected keypoints with the ones detected without additive noise. The repeatability result is shown in Fig. 3.6. Compared to the repeatability results by regular anisotropic diffusion method and isotropic diffusion method, our method is much more robust under noise. The main reason is that those two methods have instabilities when moving from finer to coarser scales as described in *Chapter 3.2.2*. They are easier to be affected by noise during the diffusion procedure. Therefore, more keypoints originally detected without noises cannot be repeatedly detected across scales under noisy circumstances.

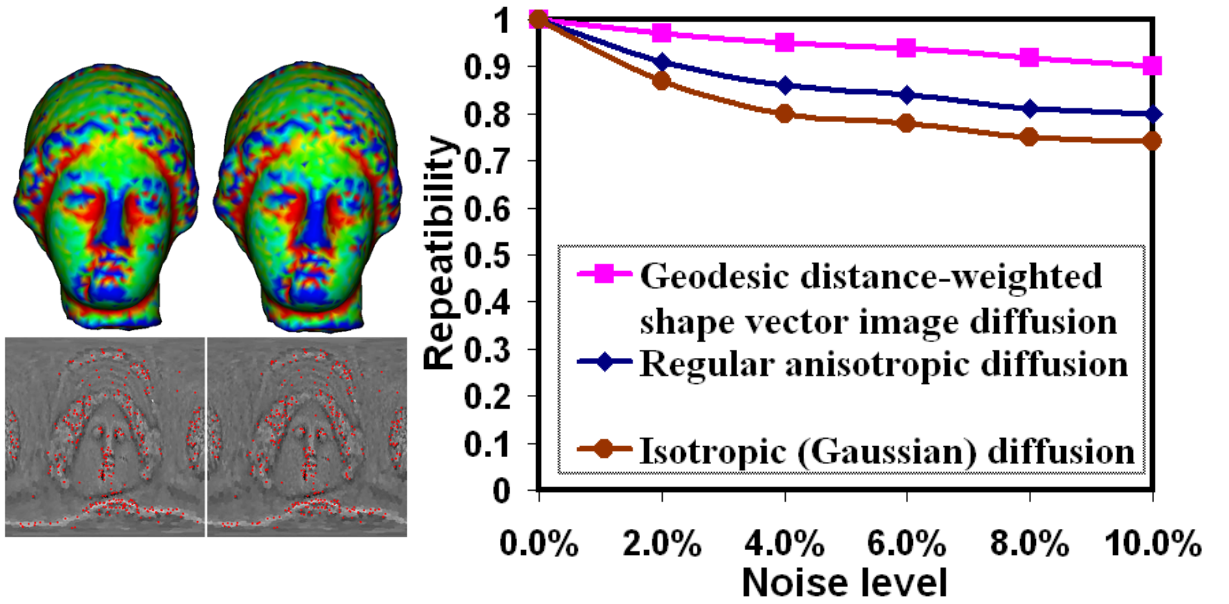


Figure 3.6: Repeatability of keypoint features when the Igea model is under different Gaussian noise levels. The left panel shows the Igea models (with the computed curvature colormaps) with 4% and 10% additive Gaussian noise and their corresponding shape vector images. The detected keypoints are shown in the shape vector images. The right panel shows the repeatability of the feature points extracted by our geodesic distance-weighted shape vector image diffusion method. The comparison to the conventional anisotropic and isotropic diffusion methods is demonstrated.

3.5.2 Surface Matching

For scanned face models, we create the shape vector images using conformal mapping. Based on this planar parameter domain, we construct the shape vector image by assigning the H and λ values to each corresponding image pixel. Hence, the shape vector image is a two dimensional vector image $[I_1, I_2]^\top$, where $I_1 = H$ and $I_2 = \lambda$. The geodesic distance is computed and encoded as well. After the shape image is generated, we use the vector diffusion to create the **DoD** matrix, of which each row is a sequence of images in different scales in each channel.

$$\begin{pmatrix} \mathbf{DoD}_1 \\ \mathbf{DoD}_2 \end{pmatrix} = \begin{pmatrix} DoD_1^{t_0} & DoD_1^{t_1} & \dots & DoD_1^{t_{n-1}} \\ DoD_2^{t_0} & DoD_2^{t_1} & \dots & DoD_2^{t_{n-1}} \end{pmatrix} \quad (3.14)$$

By finding the maxima and minima in each row of the matrix as the keypoints, the descriptor is computed for each point. Each descriptor is a $2 \times 128 = 256$ dimension vector and all these descriptors form the descriptor database. The matching algorithm is performed to find the matched points which satisfies the inequality 3.13. Fig. 3.7 shows the matching result of two faces with different expressions from the same human subject. The average matching accuracy of 10 such experiments on 10 different subjects reaches 95% in terms of correct keypoint correspondence.

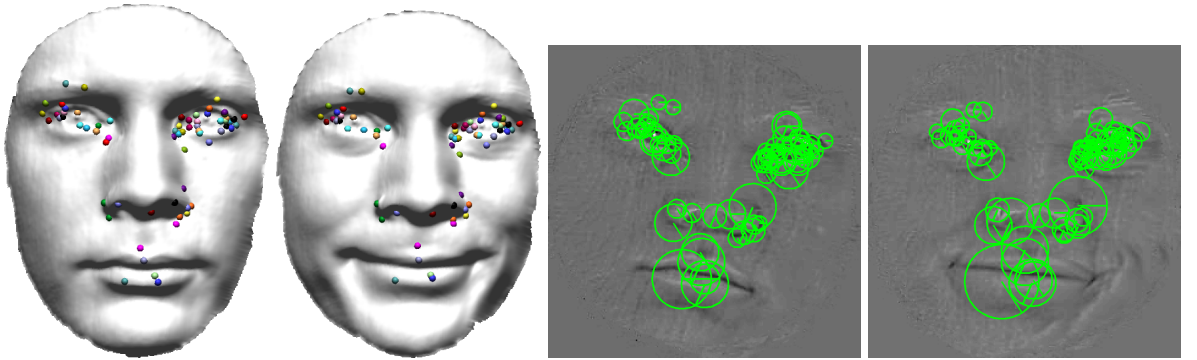


Figure 3.7: Matching of face models with different expressions from the same subject. The left panel shows all the matched keypoints between the two surfaces. The right panel shows the scales of the matched keypoints.

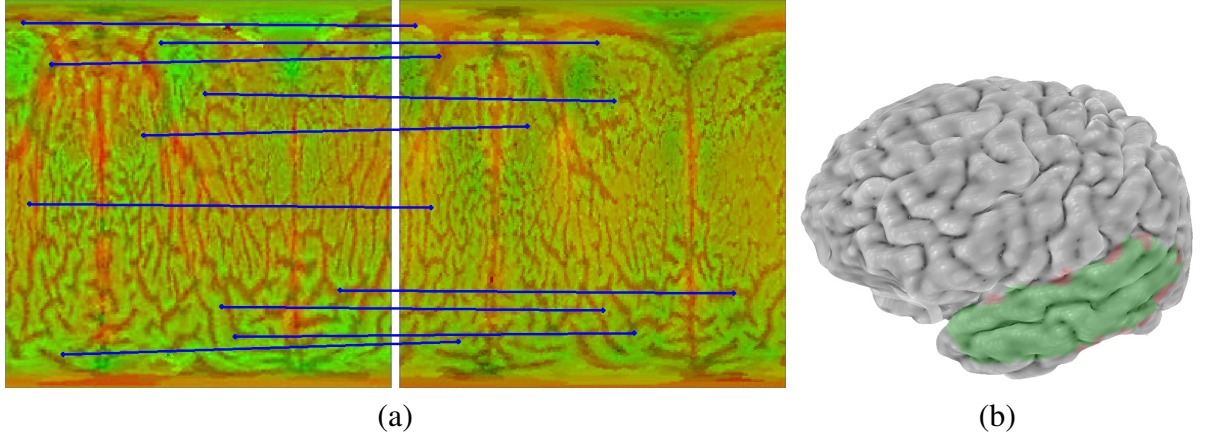


Figure 3.8: Matching of two different subjects' brain surfaces. (a) 10% of matched points are shown using the linked lines. (b) The overlap test on one registered brain region.

For the neocortex surface, a genus zero surface, conformal mapping is performed to transfer it to a sphere. The sphere can be mapped to a 2D domain through a reparameterization as follows,

$$\sigma(\theta, \varphi) = (\cos \theta \cos \varphi, \cos \theta \sin \varphi, \sin \varphi),$$

where θ and φ are the rows and columns in the 2D domain image. Then, we follow the same procedure as we use for face models to find the matching keypoints. Fig. 3.8(a) shows the matching result of two different subjects. In order to allow readers clearly see the matched points between the two shape images, only 10% of the matched points are shown in the figure. After matching, we use the matched points as landmarks to register the shape vector images using the thin-plate spline technique. We have conducted the evaluation on intersubject matching of 20 brain surfaces. The results are evaluated quantitatively in terms of major landmark (e.g., the central sulcus, the sylvian fissure, the posterior sulcus) overlaps. Fig. 3.8(b) shows one region that we used to test the registration accuracy, where the green color indicates completely correct overlap while the red color indicates mismatched areas among all the subjects. The average mismatch distance error for total 20 different subjects is only 3.98 *mm*, which outperforms the latest reported results on inter-subject brain surface registration [14].

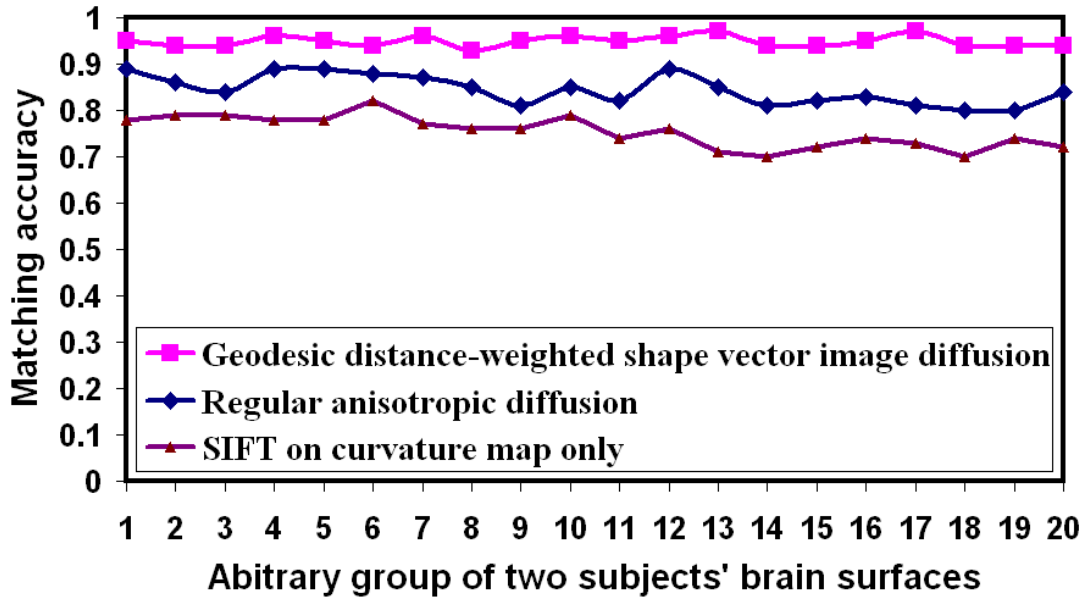


Figure 3.9: The experimental results on matching arbitrary two different brain surfaces randomly selected from 20 subjects. The comparison shows that our method constantly outperforms the regular anisotropic diffusion method and the SIFT method.

To further show the efficacy of our approach, we have compared our approach with the closely-related methods, anisotropic diffusion method and SIFT. Since SIFT can only work on scalar image, we only input the curvature channel to the SIFT processing. The regular anisotropic diffusion method is applied on both the curvature and conformal factor channels. We randomly select a pair of brain surfaces among 20 subjects. Then, our geodesic distance-weighted shape vector image diffusion method, the regular SIFT method, and the anisotropic diffusion method are performed for matching and registration. The comparison results are shown in Fig. 3.9. The main advantage of our method is to introduce geodesic distance into diffusion space. Therefore, it increases the stability of extrema detection as described in *Chapter 3.2.2* and the robustness of shape descriptors. The experimental results confirm that the keypoints and constructed local shape descriptors together are very robust features well suitable for the matching purpose. The computational time of the geodesic distance-weighted shape vector image diffusion-based feature extraction and matching is recorded for the tested

models in Table 3.1. Note that, the geodesic distance information is pre-computed offline and is not included in the recorded time.

Model	# vertices	Generate SVI (s)	SVI Size	SVI Diffusion & Matching (s)
Igea	5,002	12	256×256	43
Brain	15,102	40	512×512	172
Face	20,376	35	256×256	45

Table 3.1: Runtime of the shape vector image (SVI) construction, feature extraction and matching.

3.5.3 Multimodality Analysis

The developed framework is ideal for cross-subject analysis and visualization of multimodal brain surface properties. In order to facilitate effective analysis of medical imaging data, especially related to the human neocortical surface, a combination of noninvasive anatomical and functional imaging, such as Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI) and Positron Emission Tomography (PET), is frequently used. These modalities provide important, complementary information over the cortex regions. During the preprocessing, a brain surface can be extracted from the MRI volume data. The registration of PET and DTI volumes to the same subject’s MRI volume can be done with the mutual information registration algorithm provided in Insight Segmentation and Registration Toolkits (ITK). The registration is easy since the data is for the same subject. In order to integrate PET and MRI data, a normal fusion approach is applied in the native space of the registered MRI and PET volumes of each subject.

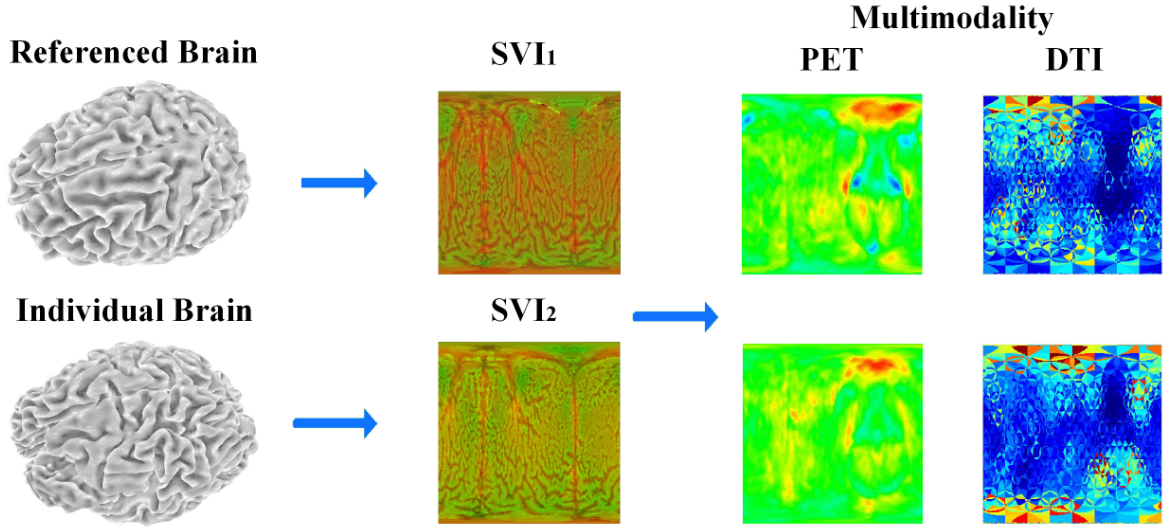


Figure 3.10: The multimodality image analysis pipeline. The referenced brain is used as the template SVI (TSVI), and then all other brain SVIs are registered based on this TSVI. Based on the registered shape vector images, multimodality data such as the PET and DTI, can be integrated over the SVI images to perform the multimodality analysis.

In this analysis, we choose one normal brain as the template in our framework. All other normal or abnormal individuals are registered to the template shape vector image (SVI) using the methods described in *Chapter 3.4*. Fig. 3.10 shows the idea and flow of our framework. The last two columns in the figure show the maps of the brain surface PET texture and DTI texture, computed from PET and DTI volumes, which are also registered across subjects because their alignments are already registered to their corresponding MRI volumes during the aforementioned preprocessing. Based on the registered SVIs, PET and DTI maps, statistical analysis of PET and DTI across subjects can be achieved.

By comparing a patient's PET texture with a group of normal subjects based upon matched SVIs, we can identify abnormal PET regions which significantly vary from the normal distribution. Fig. 3.11(a) shows two detected abnormal regions on the PET shape image. Because we know the mapping and parameterization, we can easily find out the abnormal regions in the actual brain surface. Fig. 3.11(b) shows the corresponding abnormal regions on the brain

PET data. The same scheme can be applied to the population-based DTI analysis. During the preprocessing, the cortico-cortical fiber tracts can be extracted using the brute-force fiber tracking method as shown in Fig. 3.11(c). Then, these fiber connectivity can be converted to the fiber connectivity strength ratio and plotted in the shape vector image domain to form a DTI texture image. The analysis framework can be used to detect the abnormal regions based on statistical comparison of DTI information between a patient and a group of normal subjects. Fig. 3.11(d) shows a DTI fiber connectivity strength image of an abnormal subject and the detected abnormal region as highlighted with a red contour.

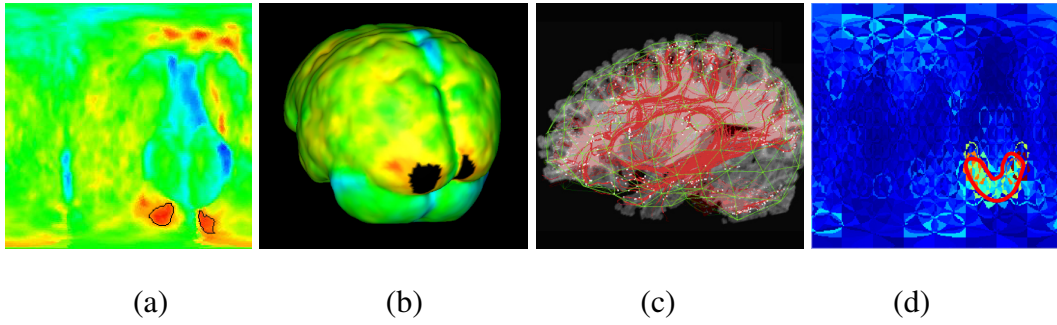


Figure 3.11: Population-based PET and DTI image analysis. (a) The regions within the black contours are the detected abnormal regions in the PET texture image; (b) The regions in black are the corresponding abnormalities on the individual’s brain cortical surface. (c) shows a 3D rendering of a normal DTI fiber connectivity. (d) shows the abnormal DTI map where the abnormality is contoured in red.

3.6 Summary

In this chapter, we have presented a novel and accurate surface matching method based on the geodesic distance-weighted diffusion of shape vector images. Through the detected stable keypoints and their local shape descriptors in the diffusion space, our method converts a 3D surface matching problem to a 2D shape vector image matching problem. The robust features facilitate the reliable matching and registration as demonstrated by our experiments. The 2D

representation allows easier statistical analysis of other modality features directly computed in the matched 2D domain. The applications to medical image analysis and visualization are demonstrated through multimodality data integration in the 2D domain to support more accurate localization of brain disorder regions using population study.

CHAPTER 4

LAPLACIAN SHAPE SPECTRUM IMPLEMENTATION AND ITS APPLICATION

4.1 Introduction and Motivation

In previous chapter, we have showed that we can geometrically map a 3D shape into a 2D vector image. Then we link the 3D shape analysis problem to the scale space processing by using the shape vector image diffusion on the 2D shape vector image. Although we can simplify a 3D problem to a 2D problem, however, the major disadvantage of the geometric mapping is that it introduces inevitable, large distortions when mapping large, complex and topologically complicated surfaces to a canonical domain. This unwanted distortion changes the characteristics of actual 3D shapes and affects the performance in shape matching or other shape analysis. In addition, it can be very difficult to compute parametrization of certain high genus-surfaces. Here we are trying to find a diffusion method which can be applied directly to the shape without the mapping.

To find the solution for shape diffusion, we have to look back to the root of 2D SIFT. Since the essential of SIFT is the diffusion equation which can create a multi-scale space, scientists are dedicating to figuring out the solution of diffusion equation for 3D shape. In [9], the spherical harmonics is used to derive the diffusion equation for a sphere from the spectral point of view. However, that only works for the sphere because it uses the spherical harmonics. It would be nice if we have the harmonics for a arbitrary regular shape instead of the spherical harmonics only for sphere. To find the harmonics, we have to turn to the shape spectrum field which provides another shape analysis tool from the spectral point of view.

Without geometric mapping from to convert the problem of 3D to 2D, we will later discuss the construction of the scale space directly on the 3D using the manifold harmonics in the

spectral point of view. And before we go to that part, in this chapter, we will discuss how to calculate the Laplacian shape spectrum and demonstrate Laplacian shape spectrum is a powerful tool for surface registration.

Specifically, in this chapter, we developed a novel method for registration between the supine and prone datasets and reduce the false positive based on their correspondence. Without matching the centerlines, we employ shape spectrum to extract the shape characteristics as the signature to find the correspondent regions between the prone and supine lumen surface. The method is simple yet efficient and accurate. Our contributions can be summarized as follows:

- A novel surface registration method based on shape spectrum is invented and applied to colon surface registration.
- We apply the algorithm to the real datasets and the experiments demonstrate the excellent accuracy of our registration results.
- We integrate the registration component into our virtual colonoscopy software for false positive reduction, which shows that it is an excellent tool aiding polyp diagnosis.

4.2 Laplacian Shape Spectrum and Manifold Harmonics

Shape spectrum is a rather new research topic in computational geometry and shape analysis [25, 29, 52, 55]. Reuter [52] defined shape spectrum as the family of eigenvalues of the Laplace-Beltrami operator on a manifold, and used it as a global shape descriptor. Lévy [29] pointed out that Laplace-Beltrami eigenfunctions are tools to understand the geometry of shapes and discussed the properties of those eigenfunctions of the Laplace-Beltrami operator. Rustamov [55] proposed a Laplace-Beltrami shape distribution based on global point signatures using the eigenfunctions and eigenvalues. Karni and Gotsman [25] used the projections of geometry on the eigenfunctions for mesh compression. The Laplace-Beltrami spectrum, which has been proved to have many good invariant properties [52], is showing tremendous power and potential in shape analysis. Most recently, Vallet and Lévy [65, 66] proposed the

manifold harmonics transform on the mesh which can be considered as a filter in the spectrum domain. These types of work inspired us to combine spectrum and diffusion together to seek a new creation of scale space for point cloud shapes from the spectrum domain by integrating the manifold harmonics with the diffusion equation which is the essential mathematical foundation for the scale space. Before we go to that far, in this chapter, we will introduce the definition of Laplacian shape spectrum and manifold harmonics.

4.2.1 Laplacian Shape Spectrum

We consider geometric attribute maps as a set of function spaces defined on a two-manifold domain. Let M denote a two-manifold, the function $f : M \rightarrow \mathbb{R}^n (n = 1, 2, \dots)$ defines a map. Different measurements of surface geometry, i.e., mean curvature, Gaussian curvature, and maximal/minimal curvature ratio, and the normals, are therefore modeled as surface maps f .

Laplace operator Δ is a second-order differential operator (the divergence of gradient) defined in Euclidean space \mathbb{R}^n . Similarly, we can define the Laplace-Beltrami operator in the n -dimensional Riemannian manifold M as the divergence of gradient

$$\Delta_M f = \frac{1}{\sqrt{|g|}} \sum_i \partial_i \left(\sum_j \sqrt{|g|} g^{ij} \partial_j f \right), \quad (4.1)$$

where g is the metric tensor over the manifold M , and g^{ij} is the element of g^{-1} . The eigen problem of Laplace-Beltrami operator can be defined as:

$$\Delta_M H = -\lambda H, \quad (4.2)$$

where λ and H are the eigen-value and corresponding eigenfunction (also called eigen-vector in discrete form). Here, the minus sign is used to ensure that all $\lambda \geq 0$.

The spectrum is defined to be the eigenvalues arranged increasingly as

$$\lambda_0 = 0 < \lambda_1 < \lambda_2 < \dots < \lambda_i < \dots \quad (4.3)$$

The appropriately normalized eigenfunction corresponding to λ_i is denoted by H_i .

The spectrum is an isometric invariant as it only depends on the gradient and divergence which in turn are defined to be dependent only on the Riemannian structure of the manifold (Eq. 4.1) [52].

4.2.2 Manifold Harmonics

Since the Laplace-Beltrami operator is Hermitian, the eigenfunctions corresponding to its different eigenvalues are orthogonal:

$$\langle H_i, H_j \rangle = \int_M H_i H_j = 0, \quad (4.4)$$

whenever $i \neq j$ and it equals 1 whenever $i = j$. The eigenfunctions H_i is called manifold harmonics and the manifold harmonics transform (MHT) is defined as following.

Given a function f on the surface, one can expand it in terms of the eigenfunctions

$$f = \tilde{f}_0 H_0 + \tilde{f}_1 H_1 + \tilde{f}_2 H_2 + \dots, \quad (4.5)$$

where the coefficients are

$$\tilde{f}_i = \langle f, H_i \rangle = \int_M f H_i. \quad (4.6)$$

Thus, the eigenfunctions of the Laplace-Beltrami operator give an orthogonal basis for the space of functions defined on the surface.

Like Fourier transform, Eq. 4.6 which calculates the coefficients is the MHT, and Eq. 4.5 is the inverse MHT [65]. And in [65], it is used to do the filtering on the shapes. In this dissertation, we will try to construct the scale space on shapes in virtue of the manifold harmonics.

4.3 Discrete Laplace-Beltrami Operator on Mesh

Eq. 4.2 can be solved by the finite element method [52]. For a triangular mesh, discrete Laplace-Beltrami operator K [39] can be applied using

$$K(p_i) = \frac{1}{2A_i} \sum_{p_j \in N_1(p_i)} (\cot \alpha_{ij} + \cot \beta_{ij})(p_i - p_j), \quad (4.7)$$

where $N_1(p_i)$ includes all the vertices, p_j , belonging to one ring neighborhood of a vertex, p_i . α_{ij} and β_{ij} are the two angles opposite to the edge in the two triangles sharing the edge i, j , and A_i is the Voronoi area of p_i .

For a triangle mesh, the Laplace-Beltrami operator can be constructed as the following matrix:

$$L_{ij} = \begin{cases} -\frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2A_i} & \text{if } i, j \text{ are adjacent,} \\ \sum_{k \in N_1(p_i)} \frac{\cot \alpha_{ik} + \cot \beta_{ik}}{2A_i} & \text{if } i=j, \\ 0 & \text{otherwise.} \end{cases} \quad (4.8)$$

where α_{ij} , β_{ij} , and A_i are the same as in Eq. 4.7 for certain i and j . Then, the spectrum problem Eq. 4.2 turns into the following eigenvalue problem:

$$L \vec{v} = \lambda \vec{v}, \quad (4.9)$$

where \vec{v} is an n -dimensional vector. Each entry of \vec{v} represents the function value at one of n vertices on the mesh. The equation above can be represented as a generalized eigenvalue problem which is much easier to solve numerically by constructing a sparse matrix W and a

diagonal matrix S such that

$$W_{ij} = \begin{cases} -\frac{\cot\alpha_{ij} + \cot\beta_{ij}}{2} & \text{if } i, j \text{ are adjacent,} \\ \sum_{k \in N_1(p_i)} \frac{\cot\alpha_{ik} + \cot\beta_{ik}}{2} & \text{if } i=j, \\ 0 & \text{otherwise,} \end{cases} \quad (4.10)$$

and $S_{ii} = A_i$. Thus, the Laplace Matrix L is decomposed as $L = S^{-1}W$ and the generalized eigenvalue problem is presented as:

$$W\vec{v} = \lambda S\vec{v}. \quad (4.11)$$

As defined above, W is symmetric and S is symmetric positive-definite. All the eigenvalues and eigenvectors are real, and the eigenvectors corresponding to different eigenvalue are orthogonal in terms of S dot-product:

$$\langle \vec{\mu}, \vec{\omega} \rangle_S = \vec{\mu}^T S \vec{\omega}, \quad (4.12)$$

where $\vec{\mu}$ and $\vec{\omega}$ are eigenvectors of Eq. 4.11.

By solving Eq. 4.11, we can get eigenvalues and eigenvectors. Fig. 4.1 shows the first four eigenvector on the colon surface with color map.

The second eigenvector of the Laplacian is called the Fiedler vector and has interesting properties, making it a good permutation vector for numerical computations [29]. Fig. 4.2 (a) and (b) show that it naturally follows the shape of the colon mesh. In other words, the Fiedler vector defines a (1-dimensional) embedding of the surface mesh. We are trying to use the embedding for the registration of our colon surface. In next two sections we will detail how we process the colon data and do the registration.

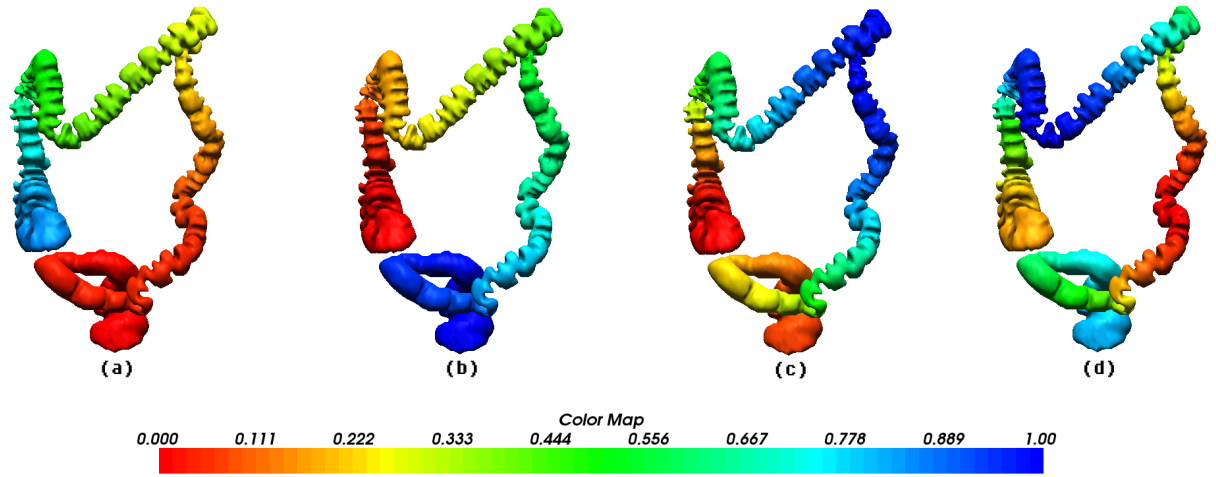


Figure 4.1: Illustration of eigenvectors with color map on the colon surface. (a) is the first eigenvector on the colon surface; (b) is the second eigenvector on the colon surface; (c) is the third eigenvector on the colon surface; (d) is the fourth eigenvector on the colon surface.

4.4 Registration Using Laplacian Shape Spectrum

The second eigenvector of the Laplacian is called the Fiedler vector and has interesting properties, making it a good permutation vector for numerical computations [29]. Fig. 4.2 (a) and (b) show that it naturally follows the shape of the mesh. In other words, the Fiedler vector defines a (1-dimensional) embedding of the surface mesh. We are trying to use the embedding for the registration of surface. First, we will introduce the colon registration problem. Next we will detail how we do the registration by using the Fiedler vector.

4.4.1 Colon Registration

Computed tomography colonography (CTC) has received increasing attention as a minimally invasive method to examination of the colon [69, 70]. CTC has shown promising results in the detection of clinically significant polyps. Using the advanced image technique, doctors can look for polyps throughout the entire colon via fly-through in a virtual colon model which is constructed from patient's abdominal images. In the last decade, many computer aided detection and diagnosis systems [4, 11] have been proposed and actively studied to improve the

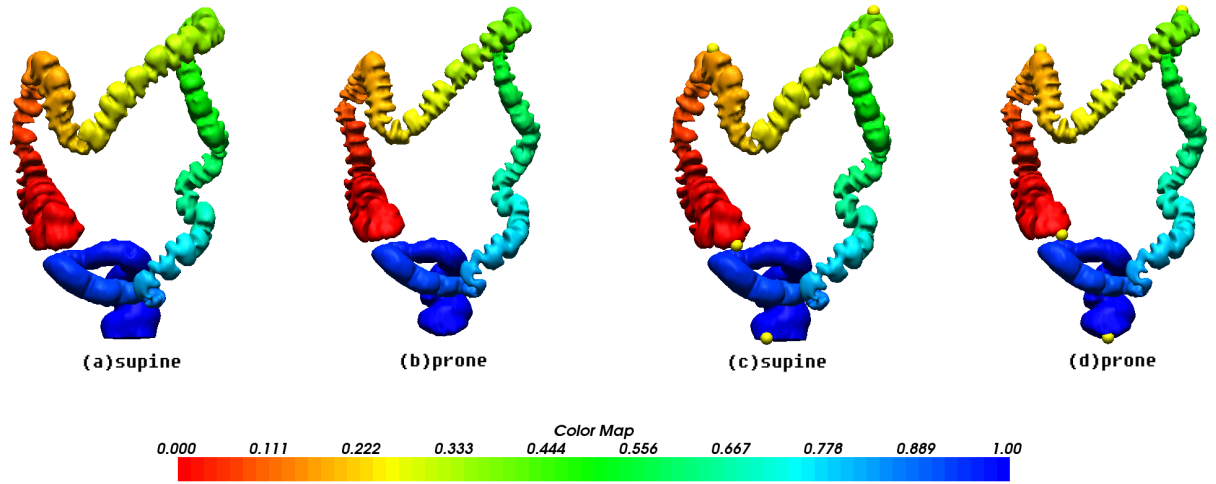


Figure 4.2: Illustration of Fiedler vector embedding and landmarks detection. The Fiedler vector gives a natural ordering of the vertex of the colon meshes in (a) and (b). The vector value has been normalized and color map. (c) and (d) show the reliable automatic detected landmarks which are rendering in yellow balls.

performance and reliability. However, most current implementation of the technique is subject to an important limitation common to all structural colon examinations: the requirement that patients undergo a rigorous purgation cleansing of the colon prior to exam. The perceived discomfort and inconvenience associated with this process has been identified as a barrier to screening [53, 79].

As a result, it has pressing demand on the electronic colon cleansing (ECC) using the image processing technology. ECC aims to remove the colonic materials from the acquired images [74]. First, the patient undergoes a less-stressful bowel preparation with oral contrast to tag the colonic materials, so that the residue stool and fluid have an enhanced image density compared with the colon/polyp tissues. Taking advantage of image segmentation and pattern recognition techniques, an ECC method can identify the enhanced colonic materials and restore a cleansed colon model for VC navigation. Several approaches are applied to the images for segmentation [79, 13, 67, 38]. Although these approaches somehow mitigate the partial volume (PV) effect that causes unexpected layers at the air-material interface and the tissue-material interface, the problem can not be perfectly solved so that false-positive (FP), namely pseudo-polyp, would

be inevitably created. Especially, sometime the small stool attachment to the colon lumen looks like polyps and it is so hard to differentiate them that it is risk to remove it during the segmentation procedure. In order to better differentiate polyps from pseudo-polyps, and to better view the lumen surface in the presence of fluid, it is common practice to obtain two CT scans of the patient, one with the patient in the prone position and one in the supine. Pseudo-polyp material may change its position between the two scans, allowing the radiologist to differentiate these structures from true polyps. Further, a second view of the colon after re-positioning may help the doctor determine if a structure is a polyp or simply a fold [44, 10].

As mentioned above, in order to better differentiate polyps from pseudo-polyps, and to better view the lumen surface in the presence of fluid, it is common practice to obtain two CT scans of the patient, one with the patient in the prone position and one in the supine. Based on this setup, the registration between the supine and prone colons is required. Several methods have been presented for registration of the supine-prone colons. Acar et al. [1] and Li et al. [31] have developed methods to map candidate polyps between the supine and prone colons using their colon wall positions relative to the colon centerlines. This kind of method only takes into account the local extrema located on the inferior/superior axis which is not reliable due to colon shifts obliquely when the patient changes position. In such a case, some local extrema may not be considered. Nain et al. [44] presented a similar approach for aligning data along the centerlines used dynamic programming. Suh and Wyatt [61] used a piece-wise centerline matching algorithm and an interpolation and extrapolation method of deformation field for deformable registration. However, the deformable model requires the good initial alignment of the two colons and strict constraints. Otherwise, it might create artifacts which, in worst case, result in pseudo-polyps. Besides the methods based on the centerlines, Näppi et al. [22] developed a directional region growing method for reducing FP based on correspondence between the supine and prone datasets. Observed from the examples they provided, the correspondences are loosely enforced. That is why only 3.7% polyp candidates was eliminated

before their Bayesian Neural Network scheme is applied. Apparently, the registration result is not good enough for diagnosis purpose.

In this chapter, we developed a novel method for registration between the supine and prone datasets and reduce the FP based on their correspondence. Without matching the centerlines, we employ shape spectrum to extract the shape characteristics as the signature to find the correspondent regions between the prone and supine lumen surface. The method is simple yet efficient and accurate. Our experiments demonstrate the excellent accuracy of our registration results.

4.4.2 Piecewise Registration

First of all, after we got the colon surface mesh, we calculated the shape spectrum of the colon mesh. By constructing matrix W and S in *Chapter 4.3*, we solved the Eq. 4.11 to get the eigenvalues and eigenvectors.

To do the registration efficiently, we first detect reliable anatomic landmarks. Based on the knowledge of colon anatomy, we expect that the mobility of the colon is at its smallest in these landmark regions. Four landmarks are established here: cecum, hepatic flexure, splenic flexure and anus. By using the normalized Fiedler vector value (F_v), the cecum and anus are detected by determination of the smallest and largest F_v . This is done both for supine and prone datasets. For hepatic flexure and splenic flexure, we first process the supine dataset then deal with the prone one. For supine dataset, the hepatic flexure is found by detecting the local maximum z-coordinate whose F_v is near 0 and the splenic flexure is found by detecting the local maximum z-coordinate whose F_v is near 1. For prone dataset, we set the neighborhood using the F_v of hepatic flexure in supine as $[F_v - \varepsilon, F_v + \varepsilon]$. At this interval in prone dataset, we detect the local maximum z-coordinate as the hepatic flexure in prone. The splenic flexure in prone is detected in the same strategy. Fig. 4.2 (c) and (d) show the landmarks both in supine and prone datasets. After this procedure, we have landmarks in sequences: $L_1^S, L_2^S, L_3^S, L_4^S$ represent the cecum, hepatic flexure, splenic flexure and anus respectively in supine dataset

and $L_1^P, L_2^P, L_3^P, L_4^P$ in prone dataset.

With the landmarks, we register the colon surface using piecewise registration. For each segment in supine between L_i^S and L_{i+1}^S , we will map it to the segment in prone between L_i^P and L_{i+1}^P , $i = 1, 2, 3$. Let

$$F_v^S(L_i^S), F_v^S(L_{i+1}^S) \quad \text{and} \quad F_v^P(L_i^P), F_v^P(L_{i+1}^P) \quad (4.13)$$

represent the Fiedler vector value at the supine and prone location of L_i^S, L_{i+1}^S and L_i^P, L_{i+1}^P respectively. Then, for a examined location L_e^S in supine between L_i^S and L_{i+1}^S , the corresponding location L_e^P in prone should have this relation of the F_v :

$$\frac{F_v^S(L_e^S) - F_v^S(L_i^S)}{F_v^S(L_{i+1}^S) - F_v^S(L_i^S)} = \frac{F_v^P(L_e^P) - F_v^P(L_i^P)}{F_v^P(L_{i+1}^P) - F_v^P(L_i^P)}. \quad (4.14)$$

It is easy to deduce that

$$F_v^P(L_e^P) = F_v^P(L_i^P) + \frac{(F_v^S(L_e^S) - F_v^S(L_i^S)) * (F_v^P(L_{i+1}^P) - F_v^P(L_i^P))}{F_v^S(L_{i+1}^S) - F_v^S(L_i^S)}. \quad (4.15)$$

Then corresponding location in prone is

$$L_e^P = (F_v^P)^{-1}(F_v^P(L_e^P)), \quad (4.16)$$

where $(F_v^P)^{-1}$ is a mapping to find the locations according to the F_v^P .

4.5 Experiments and Results

4.5.1 Data Acquisition and Pre-processing

Each patient was limited to the low-fibre diet beginning 1 day before the scheduled morning CTC. Colonic catharsis was achieved with 250ml 20% mannitol on the evening before

the examination. Before the examination, the colon was distended with 1500mL of water-soluble iodinated contrast medium ($8gI + /100ml$) using a manual insufflators with a small rectal catheter. Examinations were performed in supine positions on a 128-MDCT scanner (SOMATOM Definition AS, SIEMENS, Germany). CT technique consisted of 5.00-mm collimation, 1.375 : 1 pitch, 1-mm reconstruction interval, 120 kVp, and 50-100 mAs. Twenty cases are used for the experiment and all cases were of diagnostic quality, contain mainly fluid, but could contain feces as well. 10 patients (50%) had 20 confirmed polyps: 5 polyps were 5-10 mm, 12 polyps were 3-4 mm and 3 polyp was under 3 mm. 10 patients (50%) were normal.

In the pre-processing step, the colonic lumens are extracted from the CTC datasets by use of a level-set segmentation method [28]. Actually, any colon segmentation technique could potentially be used for the pre-processing step. Meanwhile we calculated the centerline using the distance field method [71] for later fly-through VC navigation purpose.

4.5.2 Experimental Results

Algorithms described in *Chapter 4.5.1* and *Chapter 4.4* are performed to process all the datasets. In order to evaluate the performance of the registration we proposed, a doctor experienced in the interpretation of CTC cases evaluated the 2D and 3D visualizations of the colon by use of the software we developed. Our software is implemented with C++ for the computationally intensive algorithms, OpenGL for visualization and GPU rendering for the 3D fly-through VC. The experiments are conducted on a Dell Precision Workstation T7400, which has a Xeon CPU with Quad Cores, 4GB RAM and GeForce 9800 GT video card. The tool displays the 3D VC of two intra-patient colons: supine position on the left and prone position on the right. By clicking on either 3D colon lumen, it finds the correspondence region in the other colon lumen, as well as the views of 2D CTC updated simultaneously to show the same location in both colons. A doctor can check any location he feels interested in and our system automatically updates the corresponding regions in an opposite position. The application of the method resulted in a registration accuracy of 12.6 ± 4.20 mm over 20 datasets. We compared

our method with the deformable model method [61] based on our datasets. The comparison result is shown in Table. 4.1. Our method out-performs the deformable model approach, and more importantly, our method do not create any artifacts resulting in pseudo-polyps but the deformable model would.

Method	Accuracy(mm)
Shape Spectrum Method	12.6 ± 4.20
Deformable Model	15.4 ± 6.30

Table 4.1: Results of comparison between shape spectrum method and deformable model.

We integrate our method to the software for FP reduction. First of all, the polyp candidates are selected by asking doctors to go through all the datasets to find any polyp-like protrusions. 78 polyp candidates from all the CTC datasets are presented, among which 20 polyps are true positives and 58 polyps are false positives. Since it is impossible for any registration algorithm to find exact point-to-point correspondence, in practices, it is common to set a interval, F_v , where the correspondent locations fall in. Specifically, the correspondent interval locations $Interval^P(L_e^S)$ in prone which are the correspondent locations to L_e^S in supine would be

$$Interval^P(L_e^S) = \quad (4.17)$$

$$[(F_v^P)^{-1}(F_v^P(L_e^P) - \varepsilon), (F_v^P)^{-1}(F_v^P(L_e^P) + \varepsilon)].$$

Then we check the polyps candidates in supine position and the correspondent interval locations in prone; and check the polyps candidates in prone position and the correspondent interval location in supine as well. Fig. 4.3 shows examples of FP reduction. By checking the correspondence, 48 FPs are spotted out, which is 83% reduction compared with independent processing of the datasets, with no true-positive is eliminated.

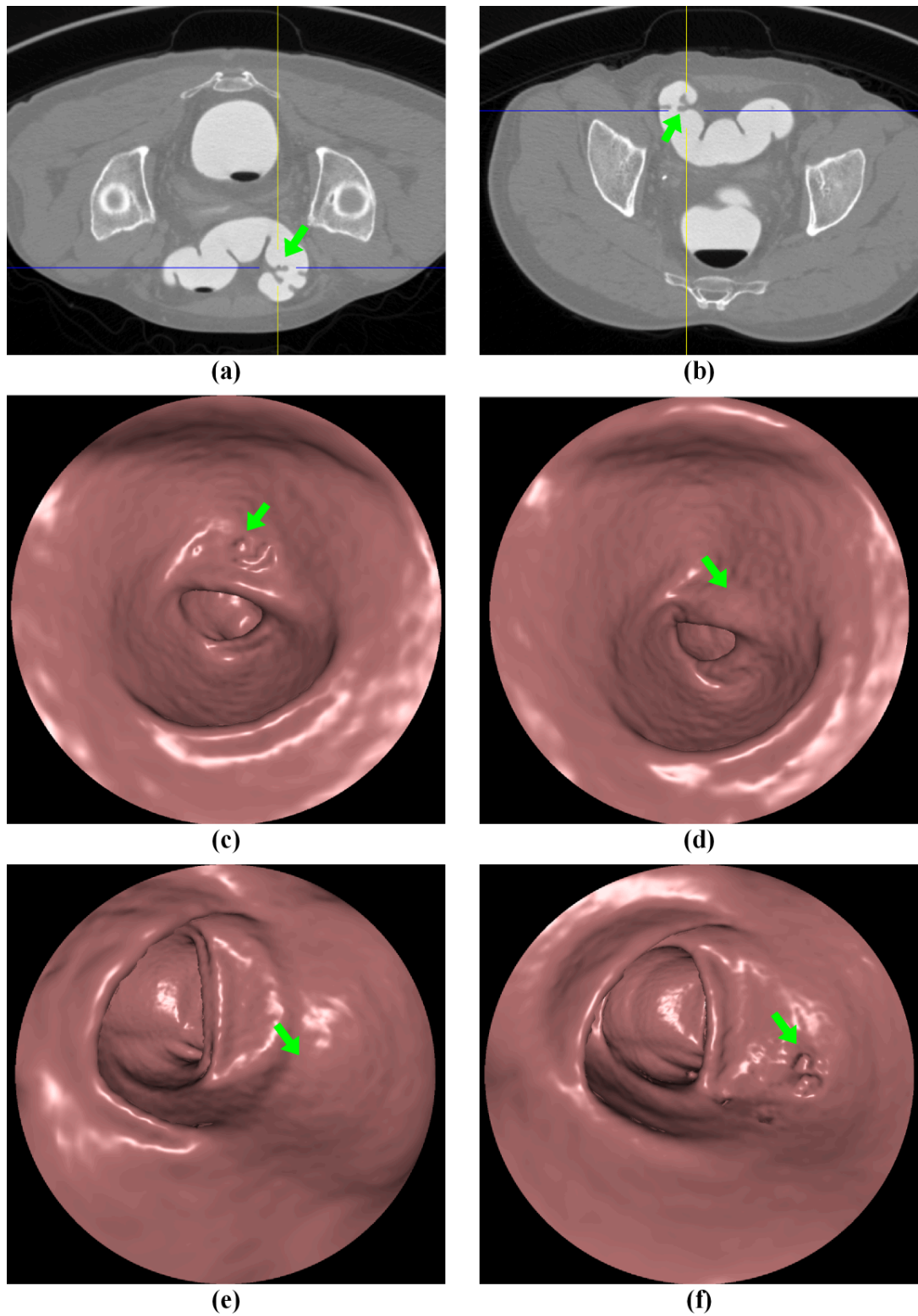


Figure 4.3: Illustration of false positive reduction. In (a), in supine position it has a false positive which the green arrow pointing at. In (b), in prone position it confirms that it is a pseudo-polyp but not a true polyp. (c) and (d) is the virtual colonoscopy of (a) and (b). (e) and (f) show another example which has false positives in prone position and can be confirmed as pseudo-polyps in supine position.

4.6 Summary

In this chapter, we show how to implement the Laplacian shape spectrum and demonstrate the surface registration using the shape spectrum. The Fiedler vector naturally following the shape makes it as a good tool for surface registration. In our test case, we have developed a piecewise method for registration of the supine and prone CTC datasets. The application of the method results in an accuracy of 12.6 ± 4.20 mm over 20 datasets with the number of false positive reduced by 83% compared with independent processing of the datasets. The experimental results indicate that the method is useful in improving the specificity of the polyps in CTC.

CHAPTER 5

SCALE SPACE CONSTRUCTION USING POINT-BASED MANIFOLD HARMONICS

In this chapter, we present a novel method to create a scale space for the point cloud from the spectral point of view by using the manifold harmonics. We rigorously derive our solution from the diffusion equation on the manifold using the point-based manifold harmonics. Utilizing the manifold harmonics, the creation of scale space on the point clouds can be achieved. Based on the multiscale structure, we can robustly detect the distinctive features by comparing the neighbors directly on the point clouds. Shape descriptors associated with the local scale information can then be defined. With these feature descriptors, the point cloud matching can be achieved through the feature descriptors matching.

5.1 Review of Related Work

With the fast development of 3D scanning and sensing technology, 3D point clouds become a more and more popular shape representation of real-world objects and scenes. There is an increasing need for methods of direct feature detection on point clouds, which can immediately propel many applications involving 3D sensing and point clouds data, such as autonomous driving by direct processing of LIDAR (Light Detection And Ranging) point clouds for road, sign and object detection, large environment scanning, simplification and visualization, feature-aware surface reconstruction from point clouds, terrain rendering, and so on. However, due to the lack of explicit connectivity and topology information in point clouds, it is much more challenging to extend many traditional, powerful methods, such as scale space processing, to this representation rather than other 3D representations (e.g., polygonal models).

In 2D image domains, scale space theory has been extensively studied and succeeded in detecting multiscale features and feature matching [33, 37, 41]. Lately, similar frameworks

have been generalized in an analogical fashion for 3D polygonal surfaces [27, 32, 47, 75, 78]. Kimmel [27] constructed a geometric scale space for images painted on a given surface. Based on surface parameterizations, scale-space shape representation and subsequent analysis can be defined on the parametric domain. For example, surface geometry is modeled as a normal map in [47]. The major disadvantage of this method is that it introduces inevitable, large distortions when mapping large, complex and topologically complicated surfaces to a canonical domain. Wu et al. [75] used a similar idea by projecting the scene to the image plane. Li and Igor [32] constructed a gaussian weighted function on the point clouds to detect multiscale features. More recently, geometric flow methods [50, 57, 82] were applied by iteratively smoothing the model to obtain a set of smoothed surfaces that constitute the scale space on shapes. This kind of methods smooth the 3D coordinates and alter the intrinsic metric which will affect the feature detection results. In general, all the methods mentioned above are just the analogous methods to scale space processing on image domains by plugging in a Gaussian kernel with different strategies for the 3D shapes. It is difficult to explicitly prove the satisfaction of the properties of the scale space theory.

Most recently, Vallet and Lévy [65] proposed the manifold harmonics transform on the mesh which can be considered as a filter in the spectrum domain analog to performing the filtering on the 2D image. Besides the work for the mesh, Belkin et al. [6] constructed Laplace operator for point clouds making it feasible to compute the eigenfunctions directly on the point cloud although the matrix form can not be guaranteed to be symmetric which is required for the manifold harmonics. All these pioneer works lead us into investigating its potential in the scale-space representation of point cloud using the manifold harmonics.

5.2 Point-based Laplace-Beltrami Operator

In this section we will detail the construction of point-based Laplace-Beltrami operator (PB-LBO) which is a symmetric operator. Finally, we will show the point-based manifold harmonics transform.

5.2.1 Computing the Approximation of PB-LBO

To build a discrete LBO converging to the continuous LBO Δ_M on the point-sampled surfaces, it is necessary to approximate $\Delta_M f(p)$ for all the sample points $p \in P$. The following is the algorithm to compute $\Delta_M f(p)$:

1. **Tangent Plane Estimation:** Set $r = 10\varepsilon$, and the point cloud P is ε -sampled [6]. Here 10ε is used to ensure that the estimated tangent plane is converging to the real tangent plane, which is proved in Belkin et al's paper [6]. Consider the point set $P_r \subseteq P$ within distance r away from p , i.e., $P_r = P \cap B(p, r)$ where $B(p, r)$ is the ball centered at p with radius r . Let Q^* be the best fitting plane passing through p such that $d(P_r, Q^*)$ is minimized. Using Har-Peled and Varadarajan's algorithm [21] (also used by Belkin et al. [6]), we construct a 2-approximation \hat{T}_p of Q^* , i.e., \hat{T}_p is a plane passing through p , and $d(P_r, \hat{T}_p) \leq 2d(P_r, Q^*)$, where $d(\cdot, \cdot)$ is the Hausdorff distance.
2. **Voronoi Cell Estimation:** Fix a positive constant $\delta \geq 10\varepsilon$, and consider the set of points P_δ that are within δ away from p , i.e., $P_\delta = P \cap B(p, \delta)$. Here $\delta \geq 10\varepsilon$ is to ensure we have enough local neighboring points for approximation. We project the points in P_δ to \hat{T}_p . When δ is sufficiently small, this projection is bijective. We denote the projection as $\hat{\Pi}$. Then we build the Voronoi Diagram of $\hat{\Pi}(P_\delta)$ on \hat{T}_p . Take the area of the Voronoi cell $V_{r_{\hat{T}_p}}(p)$ on \hat{T}_p as an approximation of the Voronoi cell area of p on surface, $V_{r_M}(p)$. $V_{r_{\hat{T}_p}}(p)$ is also denoted as $V_{r_{\hat{T}}}(p)$ here to simplify the notation. When the point cloud P gets denser, the area of $V_{r_{\hat{T}}}(p)$ is converging to the area of $V_{r_M}(p)$.
3. **Integration Approximation:** We compute $\Delta_P^t f(p)$, an approximation to $\Delta_M f(p)$, as follows:

$$\Delta_P^t f(p) = \frac{1}{4\pi t^2} \sum_{q \in P_\delta} (e^{-\frac{\|q-p\|^2}{4t}} (f(q) - f(p)) \text{vol}(V_{r_{\hat{T}}}(q))). \quad (5.1)$$

Here $\text{vol}(\cdot)$ denotes the area of the given Voronoi cell. Because the new operator is defined on the point cloud P and employs the parameter t , here we denote it as Δ_P^t .

Similar to Belkin et al's method [6], $t(\varepsilon) = \varepsilon^{\frac{1}{2+\xi}}$, and $\xi > 0$ is an arbitrary selected positive fixed number, used to ensure the convergence of Δ_P^t . It can be proved that as the points get denser, Δ_P^t will converge to Δ_M . In next section, we assemble Δ_P^t into its matrix form \hat{L}_P^t .

5.2.2 Construction of the Matrix Form for PB-LBO

Belkin et al. claimed that their discrete LBO matrix L_P^t [6] is converging point-wisely. However, their L_P^t is not guaranteed to be symmetric. In our application, to build the orthogonal Manifold Harmonic Bases, it is necessary to have a symmetric discrete Laplacian operator. A trivial way is to use $(L_P^t + L_P^{tT})/2$ instead of L_P^t . However, this trivial extension does not converge.

Next we will show how to create a symmetric operator, for more detail please refer to [35]. The LBO matrix \hat{L}_P^t can be build from the Eq. 5.1 which is linear on the function values $f(p_i)$, for $p_i \in P$. Thus it can be written as $\Delta_P^t f(p_i) = \hat{R}_i^T f$, where \hat{R}_i is an N -vector, $f = [f(p_1), f(p_2), \dots, f(p_N)]^T$ is the N -vector representing the input continuous function f sampled at the points, and $N = |P|$. Thus the matrix form \hat{L}_P^t of the discrete LBO Δ_P^t over the point cloud:

$$\Delta_P^t f = \hat{L}_P^t \cdot f, \quad (5.2)$$

where \hat{R}_i^T is the i -th row of matrix \hat{L}_P^t . Then \hat{L}_P^t can be rewritten as matrix form $\hat{L}_P^t = B^{-1} \cdot Q$, where the elements q_{ij} of the symmetric matrix Q , and the diagonal elements b_{ii} of the diagonal matrix B can be computed as follows:

$$q_{ij} = \text{vol}(V_{r_{\hat{T}}}(p_i))\text{vol}(V_{r_{\hat{T}}}(p_j)) \frac{1}{4\pi t^2} e^{-\frac{\|p_i - p_j\|^2}{4t}}, \quad (5.3)$$

$$\text{where } i \neq j, \|p_i - p_j\| \leq \delta, \quad \text{and } t(\epsilon) = \epsilon^{\frac{1}{2+\epsilon}}, \epsilon > 0,$$

$$q_{ii} = - \sum_{j \neq i} q_{ij}, \quad (5.4)$$

$$b_{ii} = \text{vol}(V_{r_T}(p_i)). \quad (5.5)$$

By redefining the functional inner product in matrix form as $\langle f, g \rangle = f^T B g$, we have:

$$\begin{aligned} \langle f, \hat{L}_P^t g \rangle &= f^T B (B^{-1} Q g) \\ &= f^T Q g \\ &= f^T Q^T (B B^{-1})^T g \\ &= (B^{-1} Q f)^T B g \\ &= \langle \hat{L}_P^t f, g \rangle, \end{aligned} \quad (5.6)$$

where means \hat{L}_P^t is a symmetric operator.

5.3 Point-based Manifold Harmonics Transform

Having the symmetric LBO matrix $\hat{L}_P^t = B^{-1} \cdot Q$, we can solve the following generalized eigen problem:

$$Q H = -\lambda B H. \quad (5.7)$$

By solving this problem, we have eigen-values $\{\lambda_i\}$ and the corresponding eigen-vectors $\{H_i\}$. $\{H_i\}$ are called the Point-based Manifold Harmonic Bases (PB-MHB) of the sampled manifold surface. PB-MHB can be used for the general spectral processing of 3D models.

All the eigenvectors are normalized, we have

$$\langle H_i, H_j \rangle = H_i^T B H_j = \delta_{ij}, \quad (5.8)$$

where

$$\begin{cases} \delta_{ij} = 0 & \text{if } i \neq j, \\ \delta_{ij} = 1 & \text{if } i = j. \end{cases} \quad (5.9)$$

and $\lambda_i \leq \lambda_j$ holds for all $i < j$. Because $\langle H_i, H_j \rangle = H_i^T B H_j = \delta_{ij}$ holds, the eigen-vectors H_i can be used to decompose functions defined over point-sample manifold surfaces:

$$\tilde{f}_i = \langle f, H_i \rangle = f^T B H_i, \quad (5.10)$$

where f is the function value sampled on the point cloud. This process is called Point-Based Manifold Harmonic Transform (PB-MHT).

5.4 Diffusion in Spectral Domain and Feature Detection

As discussed in *Chapter 5.3*, a function on the point cloud can be decomposed using the manifold harmonic by the PB-MHT. Based on that, this section describes a novel diffusion-based algorithm to extract distinctive features from the point cloud. Through the diffusion in spectral domain, we can identify the robust keypoints and their scales from the computed diffusion extrema, which are suitable for the matching purpose. In this section, firstly, we deduce the diffusion in the spectral domain so that we can create the scale space using the PB-MHT. Then we discuss how to detect the features based on the diffusion.

5.4.1 Construct Scale Space by Diffusion in Spectral Domain

It is already proved that, in a planar domain, smoothing a function with the Gaussian kernel,

$$g(x, t) = \frac{1}{4\pi kt} \exp\left(\frac{-x^2}{4kt}\right), \quad (5.11)$$

is equivalent to making the function evolve under the diffusion equation,

$$\Delta f(x, t) = \frac{1}{k} \frac{\partial}{\partial t} f(x, t), \quad (5.12)$$

where Δ is the Laplace operator. When the function f is defined on a manifold, it can be written in terms of the corresponding manifold harmonics,

$$f = \sum_i \tilde{f}_i H_i, \quad (5.13)$$

where $H_i, i = 0, 1, 2, \dots$ are the manifold harmonics, which can be computed as the eigenfunction of the manifold Laplace operator discussed in *Chapter 5.3* using Eq.5.10, and $\tilde{f}_i, i = 0, 1, 2, \dots$ are the corresponding coefficients by decomposing f with these harmonic basis functions. Here, Δ_M is the Laplace-Beltrami operator, which is a general Laplace operator defined on the manifold as mentioned in *Chapter 4.2*. When we replace the Laplace operator Δ in Eq. 5.12 with the Laplace-Beltrami operator Δ_M , given that,

$$\tilde{f}_i^t = \tilde{f}_i \exp(-\lambda_i kt), \quad (5.14)$$

it is easy to verify that its solution for Eq. 5.12 has the following form,

$$f(x, t) = \sum_i \tilde{f}_i^t H_i = \sum_i \tilde{f}_i \exp(-\lambda_i kt) H_i. \quad (5.15)$$

Let LHS denote the left-hand side and RHS denote the right-hand side of the equation, since Laplace-Beltrami operator is a linear operator according to Eq. 4.2, by plugging in Eq. 5.15, the LHS of the Eq. 5.12 becomes

$$\begin{aligned} \Delta_M f &= \sum_i \tilde{f}_i \exp(-\lambda_i kt) \Delta_M H_i \\ &= \sum_i \tilde{f}_i \exp(-\lambda_i kt) (-\lambda_i) H_i. \end{aligned} \quad (5.16)$$

The RHS of Eq. 5.12 will be

$$\begin{aligned}
\frac{1}{k} \frac{\partial f}{\partial t} &= \frac{1}{k} \sum_i \tilde{f}_i H_i \partial(\exp(-\lambda_i k t)) / \partial t \\
&= \sum_i \tilde{f}_i H_i (-\lambda_i) \exp(-\lambda_i k t) \\
&= \sum_i \tilde{f}_i \exp(-\lambda_i k t) (-\lambda_i) H_i.
\end{aligned} \tag{5.17}$$

This result shows that the diffusion of f is equivalent to the diffusion of coefficients $\tilde{f}_i, i = 0, 1, 2, \dots$ with kernel function $\exp(-\lambda_i k t)$. In other words, the effect of Eq. 5.15 is equivalent to convoluting a Gaussian functions of increasing deviations with function f on the manifold. First of all, for $t = 0$,

$$\tilde{f}_i^0 = \tilde{f}_i \exp(-\lambda_i k \cdot 0) = \tilde{f}_i. \tag{5.18}$$

The diffusion has no effect on \tilde{f}_i as expected. A typical requirement for a scale space is the *semigroup-property* [9]. That means applying diffusion to an already diffused image turns out to have the same effect as diffusing the image once, where the time-parameter is the time-parameters of the concatenated diffusions. We can easily show that our framework has this property through the following evidence:

$$\begin{aligned}
(\tilde{f}_i^s)^t &= \tilde{f}_i \exp(-\lambda_i k s) \exp(-\lambda_i k t) \\
&= \tilde{f}_i \exp(-\lambda_i k (s + t)) = \tilde{f}_i^{s+t}.
\end{aligned} \tag{5.19}$$

Under this framework, by using Eq. 5.13, Eq. 5.14 and Eq. 5.15, we can now create the scale space structure for point clouds by the diffusion on point-based manifold harmonics. For a given function f , the coefficients can be obtained by the PB-MHT. Using the decomposition equation, we calculate the \tilde{f}_i^t for some scale t . Then we can reconstruct the function $f(x, t)$ at



Figure 5.1: Diffusion on the camel point cloud model. From left to right are the diffused curvature function with t increasing. As the images show, as the t increases, the curvature function on the camel model becomes smoother and smoother.

t using Eq. 5.15. In Alg. 1, the detail of the algorithm is described regarding how to calculate the diffusion.

Algorithm 1 Diffusion using the point-based manifold harmonics

Require: the function f on point clouds

- 1: Calculate \tilde{f}_i and H_i to form the manifold harmonics using Eq. 5.13 which is to solve the eigen problem Eq. 5.7.
 - 2: Execute the diffusion using Eq. 5.14 to obtain \tilde{f}_i^t .
 - 3: Calculate the diffused function $f(x, t)$ using Eq. 5.15.
-

Then the scale space is generated:

$$\begin{pmatrix} f^{t_0} & f^{t_1} & f^{t_2} & \dots & f^{t_n} \end{pmatrix}, \quad (5.20)$$

which is a sequence of diffusion functions on the manifold, with t as the scale, in a vector format, i.e., each column element, f^{t_i} , is the diffused function at scale t_i . Fig. 5.1 shows an example of the diffusion on the camel model.

5.4.2 Keypoint Detection

In the general scale space analysis of 2D images, all the extrema are detected as keypoints through all the scales. Therefore, we propose to detect the extrema across the scale space as our keypoints since they are most robust points at the specific scales which are able to correctly represent the original point cloud. We sample the scale space by computing a sequence of diffused functions on point clouds at discrete scales, t . For each diffusion scale, we use Alg. 1

to calculate its diffused function which can be expressed in a vector form as Eq. 5.20.

In order to extract the cross-scale extrema, we compute the difference-of-diffusion (f_{DoD}) using the following equation,

$$f_{DoD}^{t_i} = f^{t_{i+1}} - f^{t_i} \quad (i = 0, \dots, n-1). \quad (5.21)$$

Consequently, we can obtain,

$$\begin{pmatrix} f_{DoD}^{t_0} & f_{DoD}^{t_1} & \dots & f_{DoD}^{t_{n-1}} \end{pmatrix}. \quad (5.22)$$

Once f_{DoD} is obtained, keypoints are identified as the local minima/maxima of the f_{DoD} across the scale domain t . Since we have to compare feature value of any point with its neighbor for finding the extrema, we have to define an neighborhood KNN for any point in the point cloud in the first place. This is achieved by using analogical geometric distance which is the embedding using shape spectrum. Specifically, given a point p on the point cloud P , we use the Global Point Signature, $GPS(p)$ [55],

$$GPS(p) = \left(\frac{1}{\sqrt{\lambda_1}} H_1(p), \frac{1}{\sqrt{\lambda_2}} H_2(p), \frac{1}{\sqrt{\lambda_3}} H_3(p), \dots \right), \quad (5.23)$$

to construct the analogical geometric distance. In [55], it was proved that Green's function can be written in terms of the eigenfunctions as

$$G(p, q) = GPS(p) \cdot GPS(q), \quad (5.24)$$

where p, q are two points in the domain. Then the dot product of two infinite-dimensional vectors shows that the inner product in the GPS domain corresponds to nothing but the Green's function. The more important thing is that the Green's function in some sense measures the extent to which two points are geometrically "bundled" together. Thus, the inner product in the

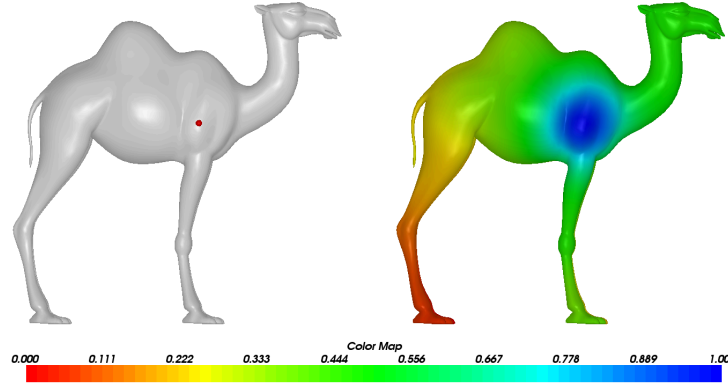


Figure 5.2: Illustration of *GPS* embedding. The left image shows the camel model and a red point – a reference point based on which we can calculate the analogical geometric distance from all other points. The right image shows the analogical geometric distance to the reference point with color map from red to blue. The bigger value, towards blue in the color map, means the closer distance between the current point and the reference point.

GPS domain is a measure of “togetherness” of two points. For more details, readers can refer to [55]. As a result, we use the *GPS* embedding to calculate the analogical geometric distance. Fig. 5.2 shows an example of the *GPS* embedding and how the analogical geometric distance works. From the example figure, it clearly shows that the *GPS* embedding can be used as the analogical geometric distance. Note that, when $G(p, q)$ is bigger, p and q becomes closer. However, for a distance function, we want it to be smaller when p and q is closer. Therefore, after normalization, we set $G(p, q) = 1 - G(p, q)$.

Based on the distance, we can define the neighborhood *KNN*. The neighborhood *KNN* of point p is extracted by k nearest points near p . *KNN* is formulated as

$$\begin{aligned} KNN &= \{\{x_1, x_2, \dots, x_k\} \mid dist(x_1, p) \leq dist(x_2, p) \\ &\leq \dots \leq dist(x_k, p), x_1, x_2, \dots, x_k \in P\}, \end{aligned} \quad (5.25)$$

where $dist(x, p)$ denotes the distance between points x and p .

With this *KNN* definition, the local extrema can be detected. It is done by comparing each point in the $f_{DoD}^{t_j}$ to its *KNN* neighbors at the same scale t_j and *KNN* neighbors in

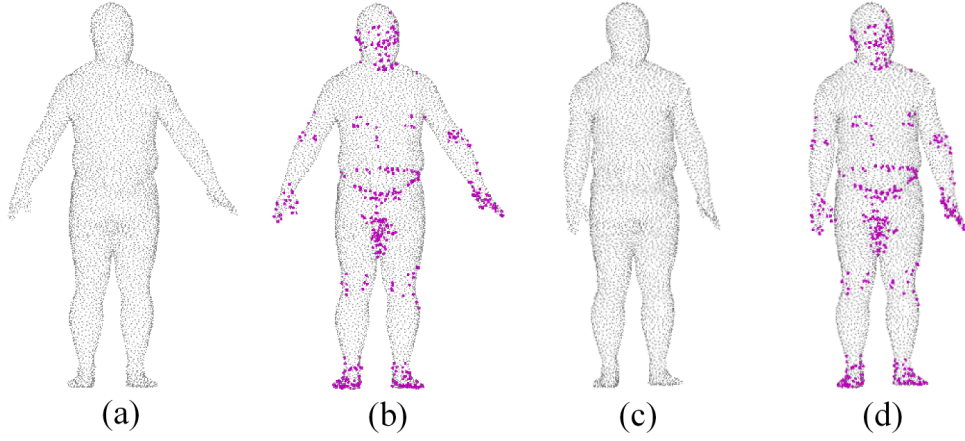


Figure 5.3: Feature point detection directly on point clouds. (a) A point cloud model; (b) The detected feature points (in red) on the point cloud model; (c) The point cloud model with a different pose; (d) The feature points (in red) directly detected from point clouds.



Figure 5.4: Keypoint detection on the camel model. From left to right are the keypoints detected on the model with scale t increasing. From the figure, it is clear that at the finer scales, more sophisticated feature points will be detected and in the coarse scales, the number of keypoints decreases.

each of the adjacent scales t_{j-1} and t_{j+1} . If the function value at the point is the maximum or minimum among all compared points, it is selected as a keypoint. This algorithm is carried out through all the f_{DoD} . The maxima and minima found in all the diffusion sub-space will be considered as the keypoints. Fig. 5.3 shows the feature point detection on the human body scan point cloud models and Fig. 5.4 shows another example of feature point detection on the camel model.

5.4.3 Curvature Function of Point Clouds

In order to apply our spectral diffusion to multiscale processing of point clouds, we need to compute the geometric attribute function f from the point clouds. Curvature-based features

are salient, with proofs from psychophysical experiments that the human visual system decomposes complex shapes into parts based on curvature. Hence, in this work, we will use curvature function instead of others function.

Assume the normal directions of the points are known, the curvature of each sample can be calculated by local fitting [76]. However, in the local PB-LBO calculation, a delaunay triangulation in the neighborhood is built, thus we can directly use this connectivity information and carry out the curvature calculation on the triangular mesh [39].

5.5 Feature Descriptor and Matching

For the purpose of matching and visualization of cross subject data, the main task is to find their local features for matching. We construct a feature descriptor for the keypoints detected using the algorithms discussed in *Chapter 5.4*. Based on the descriptors, the matching of point cloud can be achieved. In this section, firstly we discuss how to create the feature descriptor and how to match the descriptor. Then, we talk about how to match the point cloud based on the feature descriptors matching.

5.5.1 Feature Descriptor and Matching

The descriptor D_p for point p is computed using a support region, defined using a neighborhood SN based on the scale information. Like the spin image [24], firstly we compute the projection of the SN into the tangent plane of p . Then based on the tangent plane plane we create a shell-sector model to form a descriptor. Detail steps are listed below:

a) Scale-based Neighborhood SN . The neighborhood SN of point p is extracted by k nearest points near p within the ball with radius r . The value of r is a function of diffusion time t , $r = \sqrt{2t}$. Then the SN is formulated as

$$SN = \{x \mid dist(x, p) < r, x \in P\}, \quad (5.26)$$

where $dist(x, p)$ denotes the distance between points x and p .

b) Tangent Plane Projection. In *Chapter 5.2.1*, the tangent plane can be estimated. Then we calculate the projection of each point $x \in SN$ on the tangent plane. Then each point x has a projection \mathcal{P}_x on the tangent plane.

c) Shell-Sector Model Histogram. Fig. 5.5 shows an example of 3 shells and 4 sectors model. The numbers on the figure are the indices. After the tangent plane projection, the projection \mathcal{P}_x lies in a area with a specific index. We calculate the mean gradient magnitudes to form a histogram in each indexed area. For example in Fig. 5.5, we will have $b_o = 12$ bins histogram and use the indices as the orientation information. Then, the orientation histogram can be formed. Eventually, the \mathbf{D}_p is a b_o dimensional vector:

$$\mathbf{D}_p = \{D_p(1), D_p(2), \dots, D_p(b_o)\}. \quad (5.27)$$

All these histogram vectors form a feature descriptor database which will be used for matching.

d) Descriptor Matching. Based on the orientation histogram, we can computer the Euclidean distance between any two descriptors $\mathbf{D}_p, \mathbf{D}_q$. When we match two descriptors, we have to rotate the bins to calculate the minimum distance between them. Taking Fig. 5.5 as an example, there are $N_s = 4$ possible alignments between two shell-sector model; for each alignment j we calculate the similarity $DS(p, q, j), j \in N_s$ and chose the smallest value $DS(p, q)$ as the similarity of these two descriptors:

$$DS(p, q, j) = \sum_{i \in b_o} \|\mathbf{D}_p(i) - \mathbf{D}_q(i)\|, \quad (5.28)$$

$$DS(p, q) = \min\{DS(p, q, j)\}, j \in N_s. \quad (5.29)$$

5.5.2 Point Cloud Matching

In our framework, shape matching is to match the feature points in the point clouds of different objects since these points are considered as reliable representative points of point set

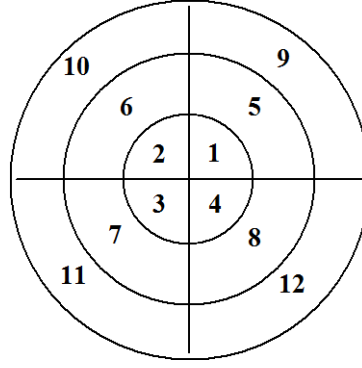


Figure 5.5: Shell-sector model: the numbers are the bin-indices.

shapes. In *Chapter 5.5.1*, we have explained how the feature descriptor for each feature point is constructed and how a feature descriptor database is formed. Next, we will discuss how to match the point clouds based on the descriptors.

We used a similar matching strategy as described in [7, 64]. Basically, it is to solve the problem of finding an optimal matching by minimizing a cost function defined over correspondences. The cost function is defined by the match quality and the geometric distortion of a correspondence. We use the similarity of these two descriptors (e.g. \mathbf{D}_i and \mathbf{D}_j) as the match cost and use the analogical geometric distance as the distortion cost. Then, an optimal correspondence can be obtained by minimizing the cost function. Specifically, a correspondence is a mapping σ indicates that p_i corresponds to $q_{\sigma(i)}$, which is sometimes abbreviated as $q_{i'}$. In other words, the cost function is defined by two terms: one is for the match quality and the other is for the geometric distortion of a correspondence:

$$cost(\sigma) = w_m Cost_{match}(\sigma) + w_d Cost_{distortion}(\sigma) \quad (5.30)$$

where the w_m and w_d are weights. The match cost for a correspondence is

$$Cost_{match}(\sigma) = \sum_i c(p_i, q_{i'}) \quad (5.31)$$

where $c(p_i, q_{i'})$ is the cost of matching p_i to $q_{i'}$ in a correspondence. We used the similarity of these two descriptors (e.g. \mathbf{D}_i and \mathbf{D}_j) as the cost, which is $DS(i, j)$ computed using Eq. 5.28 and Eq. 5.29. The distortion measure is expressed as

$$Cost_{distortion}(\sigma) = \sum_{i,j} U(p_i, q_{i'}, p_j, q_{j'}) \quad (5.32)$$

where $U(p_i, q_{i'}, p_j, q_{j'})$ is the distortion cost of mapping model points p_i to $q_{i'}$ and p_j to $q_{j'}$, respectively. We restrict these measures based on the two analogical geometric distances $r_{ij} = G(p_i, p_j)$ and $s_{i'j'} = G(q_{i'}, q_{j'})$ using Eq. 5.24. With these definitions, the distortion cost becomes

$$Cost_{distortion}(\sigma) = \sum_{i,j} d_l(\sigma), \quad (5.33)$$

where

$$d_l(\sigma) = \frac{|s_{i'j'}| - |r_{ij}|}{|r_{ij}|}. \quad (5.34)$$

In order to find an assignment to minimize the cost function described by the terms in Eq. 5.31 and Eq. 5.32 above, the correspondence problem can be written as an Integer Quadratic Programming (IQP) problem:

$$cost(x) = \sum_{a,b} U(a, b) x_a x_b + \sum_a c(a) x_a \quad (5.35)$$

where the binary indicator variable x has entries x_a , that if 1, indicate $\sigma(a_i) = a_j$; $c(a) = c(a_i, a_j)$ and $U(a, b) = U(a_i, a_j, b_i, b_j)$ from Eq. 5.31 and Eq. 5.32. We constrain x to represent an assignment, $\sum_j x_{ij} = 1$ for each i , by writing x_{ij} in place of $x_{a_i a_j}$. These linear constraints are encoded as A and b . Eventually, we are going to solve the IQP in matrix form as:

$$\begin{aligned} mincost(x) &= x' U x + c' x \quad \text{subject to,} \\ Ax &= b, \quad x \in \{0, 1\}^n \end{aligned} \quad (5.36)$$

By solve Eq. 5.36, we can obtain the optimal correspondences. For more details about solving this equation, readers can refer to [7, 64].

5.5.3 Experiments and Results

Our algorithm is implemented with C++ for the computationally intensive algorithms and VTK/OpenGL for rendering and visualization. The experiments are conducted on a Dell XPS 630i, which has Intel Core(TM)2 Quad CPU Q9550 with 4GB of RAM.

3D shape matching is a fundamental task in computer vision and geometry processing. This experiment is performed on isometric matching and non-isometric matching to analyze the performance of our proposed framework. A scale-associated shape representation is employed as the local shape descriptor and the descriptor matching is obtained using the algorithm discussed in *Chapter 5.5*. Firstly, the method is evaluated under isometric transformation. Secondly, we test our framework on a real face scans database which has different expressions of human subjects.

a) Isometric matching

This experiment is to test the feature detection and matching under isometric transformation. The input point cloud is a scanned human model shown in the left image in Fig. 5.6. Then, the character put the hands down to generate another scan shown in the right image in Fig. 5.6. We perform our feature detection algorithm and point cloud matching algorithm to obtain the matching as shown in Fig. 5.6, where the lines linked between feature points indicate the correspondence. The matching result shows that the algorithm works well for matching objects with isometric deformations.

b) Face matching

This experiment is performed on the 3D face database to analyze the performance of our proposed framework. Two types of experiments are conducted: (1) matching the same subject with different expressions (i.e., undergoing non-isometric deformation); (2) matching different subjects. Our method can correctly match the same subject with different expressions while

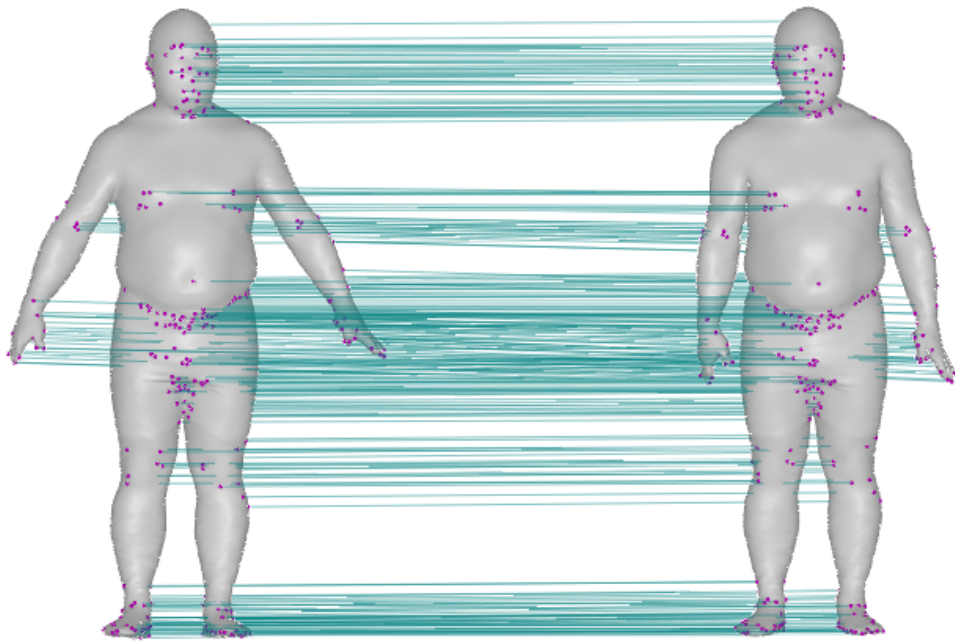


Figure 5.6: Illustration of the matching result. Matched points are connected by linked lines.

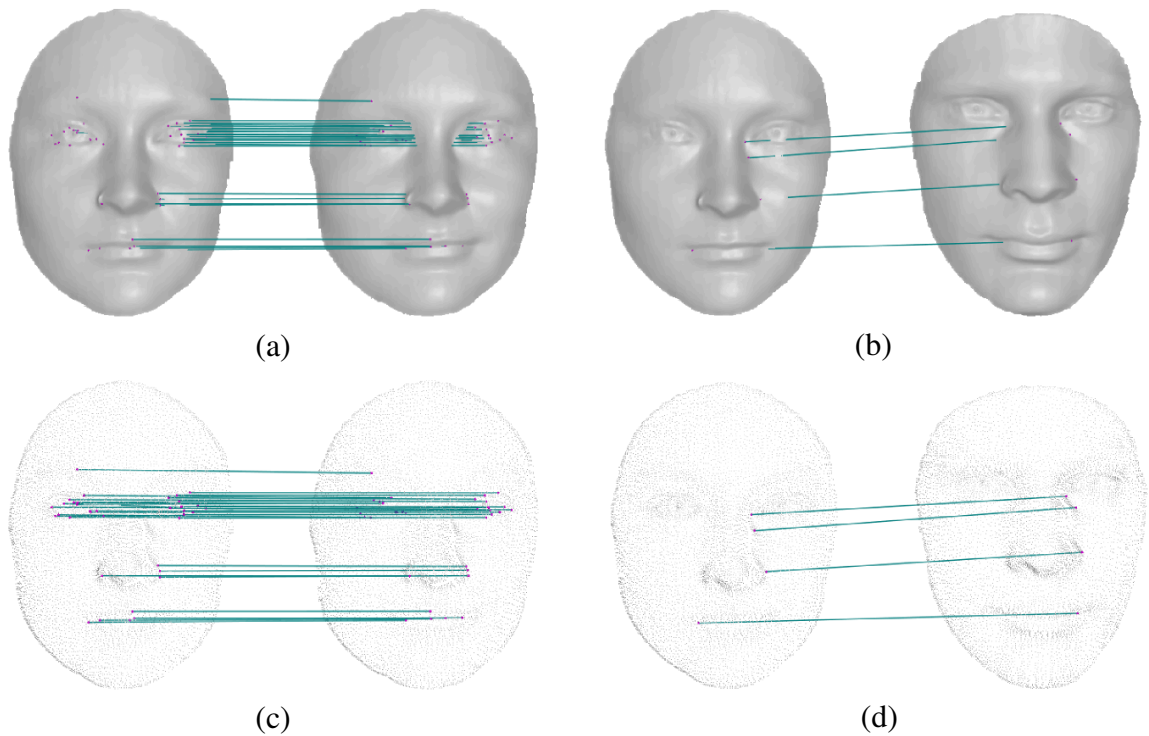


Figure 5.7: Face matching. (a) Matching between the same subject with different expressions. (b) Matching between two different subjects. (c) and (d) are the point cloud views of (a) and (b) respectively.

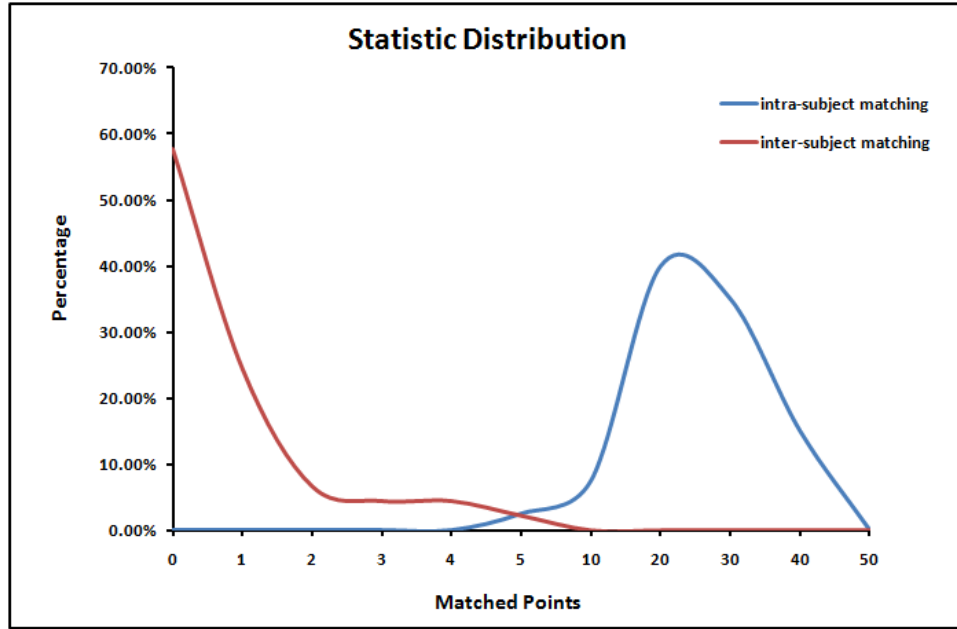


Figure 5.8: Statistical results. The red curve shows the inter-subject matching and the blue one shows the intra-subject matching.

differentiating different subjects based on the number of matched keypoints.

The example of face matching is shown in Fig. 5.7. The left image in Fig. 5.7(a) is the neutral expression and the right image is the smile expression. In Fig. 5.7(b), it shows two faces from two different subjects. The result shows the differentiation between two different subjects, where the number of matched keypoints is significantly fewer than the case in Fig. 5.7(a). Then, we conducted a statistical experiment on 20 subjects with 5 different expressions each subject. Fig. 5.8 shows the statistics of the numbers of matched keypoints among the 20 subjects. As we can see, the numbers of matched keypoints are significantly greater in the intra-subject matching (same subject different expressions) against those found in the inter-subject case (between different subjects). The number of matched keypoints is descriptive enough to differentiate models from different subjects and meanwhile retrieve distinct expressions of the same person. The experiments indicate that our method is effective for the face point cloud matching and has great potential for the retrieval task in face scan data.

5.6 Summary

In this chapter, we have presented a framework to generate the scale space on point clouds using point-based manifold harmonics. Built upon the point-based manifold harmonics transform, we generalize the diffusion function directly on the point clouds to create the scale space. Subsequently, feature detection can be performed to detect keypoints together with their scale size. We construct a multiscale descriptor for each feature point which is effective for matching point clouds. The performance of our framework is thoroughly evaluated through our experiments which demonstrate excellent matching performance for point cloud data.

CHAPTER 6

CONCLUSION

The purpose of this chapter is twofold. We first summarize the contributions made by this dissertation, then we point towards the future work.

6.1 Contributions

Shape analysis is a fundamental aspect of many problems in computer graphics and computer vision, including surface matching and surface registration. There has been a lot of research on 3D shape analysis. Modern methods are trying to imitating the 2D SIFT to construct a multiscale structure for 3D shape. Our research focus on developing a multiscale structure in shape vector image using the geometric mapping and the one directly on the 3D shape using manifold harmonics. Our main contributions include:

- *Shape Vector Image Diffusion Framework (Chapter 3).* We try to link the 3D shape analysis to the scale space processing by using the geometric mapping. In virtue of the geometric mapping, we construct the shape vector image for the 3D shape. We then propose the shape vector image diffusion framework for multiscale diffusion space construction. In the diffusion space, we detect the cross-scale extrema as the feature points. Then we construct feature descriptors for feature matching. The matching of feature points can be utilized for shape matching and registration. Experiments demonstrate the excellent performance of the proposed method and it is a powerful tool for analyzing surfaces with multidimensional textures.
- *Laplacian Shape Spectrum Implementation and Its Application (Chapter 4).* We introduce Laplacian shape spectrum and implement the Laplace-Beltrami operator on mesh using the discrete settings. By investigating the Laplacian shape spectrum, we propose the surface registration using the Fiedler vector from Laplacian shape spectrum. We reg-

ister the colon surface using piecewise registration using the Fiedler vector value. We apply the algorithm to the real CTC datasets and our experiments show excellent registration results on colon surface registration and it shows excellent potential in reducing the false positive when it is used to determine polyps through correspondences between prone and supine images.

- *Scale Space Construction Using Point-based Manifold Harmonics (Chapter 5)*. In order to construct the scale space for 3D shape without using the geometric mapping, we rigorously derive our solution from the diffusion equation on the manifold using the manifold harmonics. Not only confined on mesh, we implement the Laplace-Beltrami operator on 3D point cloud data. By constructing the scale space directly on point cloud, based on the multiscale structure, we can robustly detect the distinctive features and construct descriptors from the neighbors using the GPS. Then the point cloud matching can be achieved through the feature descriptors matching. Our experiments demonstrate that our method is showing excellent matching performance for point cloud data.

These contributions are described in more detail and validated in the main body of the dissertation. Please refer to the corresponding chapters for details.

6.2 Future Work

This dissertation work also opens several venues for future work, with the focus on visual object processing and analysis.

1. 2D/3D feature matching. We try to propose a novel method to match the 2D/3D features automatically. As discussed in previous chapters, to extract the 3D feature points, we can use the manifold harmonics to create the scale space in virtue of the shape spectrum for the 3D feature detection. After the feature detection, for each feature point we adopt perspective projection to project the characteristics to a image. Then we will try SIFT matching on the 2D texture image and perspective projected images. Finally, through the

relationship between the perspective projected images and 3D object, we can identify the correspondences between the 2D image and the 3D object.

2. Shape retrieval. The topic has recently gained popularity in computer vision and pattern recognition communities. With the development of 3D scanning technique, large databases of 3D models available in the public domain have created the demand for shape search and retrieval algorithms capable of finding similar shapes in the same way a search engine responds to text or image queries. Since many shapes manifest rich variability, shape retrieval is often required to be invariant to different classes of transformations and shape variations. One of the most challenging settings is the case of deformable shapes, in which the transformations vary in different forms. I plan to propose a shape spectrum based invariant characteristic for the deformable shape. I will explore its shape spectrum features and try to combine with other statistical methods. The long-term goal in this research direction embodies two lines: retrieval of different actions of the same object and retrieval of the same action of different objects.

APPENDIX

Publications

• *Journals*

- [1] V. Taimouri, X. Liu, **Z. Lai**, C. Liu, D. Pai, and J. Hua, “Colon Segmentation for Prepless Virtual Colonoscopy”, *IEEE Transactions on Information Technology in Biomedicine*, 2011 (To appear).
- [2] G. Zou, J. Hua, **Z. Lai**, X. Gu, and M. Dong, “Intrinsic Geometric Scale Space by Shape Diffusion”, *IEEE Transactions on Visualization and Computer Graphics (VIS)*, Vol. 15, No. 6, pp. 1193 - 1200, 2009.
- [3] C. Lin, S. Lu, X. Fei, A. Chebotko, D. Pai, **Z. Lai**, F. Fotouhi, and J. Hua. “A Reference Architecture for Visual Scientific Workflow Management Systems and the VIEW SOA Solution”, *IEEE Transactions on Service Computing*, Vol. 2, No. 1, pp. 79-92, 2009.
- [4] J. Hua, **Z. Lai**, M. Dong, X. Gu, and H. Qin, “Geodesic Distance-Weighted Shape Vector Image Diffusion”, *IEEE Transactions on Visualization and Computer Graphics (VIS)*, Vol. 14, No. 6, pp. 1643 - 1650, 2008.

• *Conferences*

- [5] **Z. Lai**, J. Hu, C. Liu, V. Taimouri, D. Pai, J. Zhu, J. Xu, and J. Hua, “Intra-patient Supine-Prone Colon Registration in CT Colonography Using Shape Spectrum”, In *Proceedings of the 13th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2010 : 332-339.
- [6] C. Liu, **Z. Lai**, J. Hu, and J. Hua, “Detail Preserving 3D Motion Compression Based on Local Transformation”, In *Proceedings of the Fourth Pacific-Rim Symposium on Image and Video Technology (PSIVT)*, 2010 : 507-514.

[7] **Z. Lai** and J. Hua, “3D Surface Matching and Registration through Shape Images”, In *Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2008 : 44-51.

[8] C. Lin, S. Lu, **Z. Lai**, A. Chebotko, X. Fei, J. Hua and F. Fotouhi, “Service-Oriented Architecture for VIEW: a VISual sciEntific Workflow Management System”, In *Proceedings of the 5th IEEE International Conference on Services Computing (SCC)*, 2008 : 335-342.

[9] A. Chebotko, C. Lin, X. Fei, **Z. Lai**, S. Lu, J. Hua, and F. Fotouhi, “VIEW: a VISual sciEntific Workflow management system”, In *Proceedings of the first IEEE International Workshop on Scientific Workflows (SWF)*, 2007 : 207-208.

- ***Papers in Preparation***

[10] **Z. Lai**, C. Liu, and J. Hua, “Point-based Manifold Harmonics for Multi-scale Feature Detection on Point Clouds”, planned submission to *IEEE Transactions on Visualization and Computer Graphics*.

[11] **Z. Lai**, C. Liu, J. Hu, V. Taimouri, D. Pai, and J. Hua, “Intra-patient Supine-Prone Colon Registration in CT Colonography Using Laplacian Shape Spectrum”, planned submission to *IEEE Transactions on Biomedical Engineering*.

BIBLIOGRAPHY

- [1] ACAR, B., NAPEL, S., PAIK, D., LI, P., J.YEE, JEFFREY, R., AND BEAULIEU, C. Medial axis registration of supine and prone CT colonography data. In *Proceedings of the 23rd Annual EMBS International Conference* (2001), pp. 2433–2436.
- [2] ALAA, AND FARAG, A. CSIFT: A SIFT descriptor with color invariant characteristics. In *Proc. CVPR* (2006), vol. 2, pp. 1978–1983.
- [3] ATHITSOS, V., ALON, J., SCLAROFF, S., AND KOLLIOS, G. Boostmap: a method for efficient approximate similarity rankings. In *Proc. CVPR* (2004), pp. pages II: 268–275.
- [4] BAKER, M. E., BOGONI, L., OBUCHOWSKI, N. A., DASS, C., KENDZIERSKI, R. M., REMER, E. M., EINSTEIN, D. M., CATHIER, P., JEREBKO, A., LAKARE, S., BLUM, A., CAROLINE, D. F., AND MACARI, M. Computer-aided detection of colorectal polyps: Can it improve sensitivity of less-experienced readers? preliminary findings. *Radiology* 245 (2007), 140–149.
- [5] BASRI, R., AND JACOBS, D. W. Recognition using region correspondences. *International Journal of Computer Vision* 25, 2 (1996), 141–162.
- [6] BELKIN, M., SUN, J., AND WANG, Y. Constructing Laplace operator from point clouds in R^d . In *Proceedings of the 19th annual ACM-SIAM Symposium on Discrete Algorithms* (2009), pp. 1031–1040.
- [7] BERG, A. C., BERG, T. L., AND MALIK, J. Shape matching and object recognition using low distortion correspondence. In *Proc. CVPR* (2005), pp. 26–33.
- [8] BROWN, M., AND LOWE, D. Invariant features from interest point groups. In *British Machine Vision Conference* (2002), pp. 656–665.

- [9] BÜLOW, T. Spherical diffusion for 3D surface smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (2004), 1650–1654.
- [10] CHEN, S., LU, D., HECHT, J., AND KADELL, B. CT colonography: value of scanning in both the supine and prone positions. *American Journal of Roentgenology* 172 (1999), 595–599.
- [11] CHOWDHURY, T., WHELAN, P., AND GHITA, O. A fully automatic CAD-CTC system based on curvature analysis for standard and low dose CT data. *IEEE Trans. Biomedical Engineering* 55 (2008), 888–901.
- [12] CHUA, C., AND JARVIS, R. 3D free-form surface registration and object recognition. In *International Journal of Computer Vision* (1996), vol. 17, pp. 77–99.
- [13] DIJKERS, J. J., VAN WIJK, C., VOS, F., FLORIE, J., NIO, Y. C., VENEMA, H. W., TRUYEN, R., AND VAN VLIET, L. J. Segmentation and size measurement of polyps in CT colonography. In *MICCAI* (2005), pp. 712–719.
- [14] ECKSTEIN, I., JOSHI, A., KUO, C., LEAHY, R., AND DESBRUN, M. Generalized surface flows for deformable registration and cortical matching. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention* (2007), vol. 1, pp. 692–697.
- [15] FLOATER, M. S., AND HORMANN, K. Surface parameterization: a tutorial and survey. *Advances in Multiresolution for Geometric Modelling* (Springer, 2004), 157–186.
- [16] FROME, A., HUBER, D., KOLLURI, R., BULOW, T., AND MALIK, J. Recognizing objects in range data using regional point descriptors. In *ECCV04* (May 2004).
- [17] FUNKHOUSER, T., P. MIN, M. K., CHEN, J., HALDERMAN, A., DOBKIN, D., AND JACOBS, D. A search engine for 3D models. *ACM Transactions on Graphics* (2003), 83–105.

- [18] GRIMSON, E., AND LOZANO-PÉREZ, T. Localizing overlapping parts by searching the interpretation tree. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9 (1987), 469–482.
- [19] GU, X., WANG, Y., CHAN, T., THOMPSON, P., AND YAU, S. T. Genus zero surface conformal mapping and its application to brain surface mappings. *IEEE Transactions on Medical Imaging* 23, 8 (2004), 949–958.
- [20] GU, X., WANG, Y., AND YAU, S. Geometric compression using riemann surface structure. *Communications in Information and Systems* 3, 3 (2004), 171–182.
- [21] HAR-PELED, S., AND VARADARAJAN, K. Projective clustering in high dimensions using core-sets. In *Proceedings of the 18th annual Symposium on Computational Geometry* (2002), pp. 312–318.
- [22] J.NÄPPI, A.OKAMURA, H.FRIMMEL, A.DACHMAN, AND H.YOSHIDA. Region-based supine-prone correspondence for the reduction of false-positive CAD polyp candidates in CT colonography. *Academic Radiology* 12, 6 (2005), 695–707.
- [23] JOHNSON, A., AND HEBERT, M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 (1999), 433–449.
- [24] JOHNSON, A., AND HEBERT, M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 1 (1999), 433 – 449.
- [25] KARNI, Z., AND GOTSMAN, C. Compression of soft-body animation sequences. *Computers & Graphics* 28 (2004), 25–34.

- [26] KAZHDAN, M., FUNKHOUSER, T., , AND RUSINKIEWICZ, S. Rotation invariant spherical harmonic representation of 3D shape descriptors. In *Eurographics/ACM SIGGRAPH symposium on Geometry processing* (2003), pp. 156–164.
- [27] KIMMEL, R. Intrinsic scale space for images on surfaces: The geodesic curvature flow. In *Graphical Models and Image Processing* (1997), pp. 365–372.
- [28] LANKTON, S., AND TANNENBAUM, A. Localizing region-based active contours. *IEEE Transactions on Image Processing* 17, 11 (2008), 2029–2039.
- [29] LÉVY, B. Laplace-Beltrami eigen functions: Towards an algorithm that understands geometry. In *SMI '06: Proceedings of the IEEE International Conference on Shape Modeling and Applications 2006* (2006), p. 13.
- [30] LÉVY, B., PETITJEAN, S., RAY, N., AND MAILLOT, J. Least squares conformal maps for automatic texture atlas generation. In *SIGGRAPH* (2002), pp. 362–371.
- [31] LI, P., NAPEL, S., ACAR, B., PAIK, D., AND BEAULIEU, R. J. J. C. Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: Pilot study. *MED PHYS* 31, 10 (2004), 2912–2923.
- [32] LI, X., AND GUSKOV, I. Multi-scale features for approximate alignment of point-based surfaces. In *SGP '05: Proceedings of the third Eurographics symposium on Geometry processing* (2005), pp. 217–226.
- [33] LINDEBERG, T. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1993.
- [34] LINDEBERG, T. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics* (1994), 224–270.

- [35] LIU, Y., PRABHAKARAN, B., AND GUO, X. Point-based manifold harmonics. Tech. rep., University of Texas at Dallas, 2009.
- [36] LOWE, D. G. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence* 31, 3 (1987), 355–395.
- [37] LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 2 (2004), 91–110.
- [38] LU, L., WOLF, M., LIANG, J., DUNDAR, M., BI, J., AND SALGANICOFF, M. A two-level approach towards semantic colon segmentation: Removing extra-colonic findings. In *MICCAI* (2009), pp. 1009–1016.
- [39] MEYER, M., DESBRUN, M., SCHRÖDER, P., AND BARR, A. Discrete differential geometry operators for triangulated 2-manifolds. In *Proceedings of the VisMath Conference* (2002), pp. 1–26.
- [40] MIKOLAJCZYK, K., AND SCHMID, C. A performance evaluation of local descriptors. In *Proc. CVPR* (2002), vol. 2, pp. 257–263.
- [41] MIKOLAJCZYK, K., AND SCHMID, C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 27, 10 (2005), 1615–1630.
- [42] MORI, G., BELONGIE, S., AND J.MALIK. Efficient shape matching using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11) (2005), 1832–1837.
- [43] MURASE, H., AND K.NAYAR, S. Visual learning and recognition of 3D objects from appearance. *International Journal of Computer Vision* 14, 1 (1995), 5–24.

- [44] NAIN, D., HAKER, S., GRIMSON, W. E. L., JR., E. C., WELLS, W. M., JI, H., KIKINIS, R., AND WESTIN, C.-F. Intra-patient prone to supine colon registration for synchronized colonoscopy. In *MICCAI* (2002), pp. 573–580.
- [45] NELSON, R. C., AND SELINGER, A. Large-scale tests of a keyed, appearance-based 3D object recognition system. *Vision Research* 38, 15 (2004), 2469–88.
- [46] NIRENBARG, L. A strong maximum principle for parabolic equations. *Communications on Pure and Applied Mathematics* 6, 2 (1953), 167–177.
- [47] NOVATNACK, J., AND NISHINO, K. Scale-dependent 3D geometric features. In *Proc. ICCV* (2007), pp. 1–8.
- [48] OHBA, K., AND IKEUCHI, K. Detectability, uniqueness, and reliability of eigen windows for stable verification of partially occluded objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 9 (1997), 1043–48.
- [49] OSADA, R., FUNKHOUSER, T., CHAZELLE, B., AND DOBKIN, D. Shape distributions. *ACM Transactions on Graphics* 22 (2002), 807–832.
- [50] PAULY, M., KOBELT, L. P., AND GROSS, M. Point-based multiscale surface representation. *ACM Trans. Graph.* 25, 2 (2006), 177–193.
- [51] PERONA, P., AND MALIK, J. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence* 12, 7 (1990), 629–639.
- [52] REUTER, M., WOLTER, F., AND PEINECKE, N. Laplace-Beltrami spectra as “shape-DNA” of surfaces and solids. *Computer Aided Design* 38, 4 (2006), 342–366.
- [53] RISTVEDT, S. L., MCFARLAND, E. G., WEINSTOCK, L. B., AND THYSSEN, E. P. Patient preferences for CT colonography, conventional colonoscopy, and bowel prepara-

- tionpatient preferences for CT colonography. *The American Journal of Gastroenterology* 98 (2003), 578–585.
- [54] RUIZ-CORREA, S., SHAPIRO, L., AND MEILA, M. A new paradigm for recognizing 3D object shapes from range data. In *ICCV* (2003), pp. 1126–1133.
- [55] RUSTAMOV, R. M. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *SGP '07: Proceedings of the fifth Eurographics symposium on Geometry processing* (2007), pp. 225–233.
- [56] SCHIELE, B., AND JAMES, L. Object recognition using multidimensional receptive field histogram. In *Fourth European Conference on Computer Vision, Cambridge, UK* (1996), pp. 610–619.
- [57] SCHLATTMANN, M., DEGENER, P., AND KLEIN, R. Scale space based feature point detection on surfaces. *Journal of WSCG* 16, 1-3 (2008).
- [58] SCHMID, C., AND MOHR, R. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 5 (1997), 530–534.
- [59] SCHOEN, R., AND YAU, S. T. *Lectures on Harmonic Maps*. International Press, Harvard University, Cambridge MA, 1997.
- [60] SHARON, E., AND MUMFORD, D. 2D-shape analysis using conformal mapping. In *Proc. CVPR* (2004), pp. II: 350–357.
- [61] SUH, W., AND WYATT, C. Deformable registration of supine and prone colons using centerline analysis. In *4th IEEE International Symposium on Biomedical Imaging* (2007), pp. 708–711.
- [62] SUN, Y., AND ABIDI, M. Surface matching by 3D points fingerprint. In *ICCV* (2001), pp. II: 263–269.

- [63] SWAIN, M., AND BALLARD, D. Color indexing. *International Journal of Computer Vision* 7, 1 (1991), 11–32.
- [64] TORRESANI, L., KOLMOGOROV, V., AND ROTHER, C. Feature correspondence via graph matching: Models and global optimization. In *ECCV '08: Proceedings of the 10th European Conference on Computer Vision* (2008), pp. 596–609.
- [65] VALLET, B., AND LÉVY, B. Manifold harmonics. Tech. rep., INRIA - ALICE Project Team, 2007.
- [66] VALLET, B., AND LÉVY, B. Spectral geometry processing with manifold harmonics. *Computer Graphics Forum (Proceedings Eurographics)* (2008).
- [67] VAN WIJK, C., VAN RAVESTEIJN, V. F., VOS, F., TRUYEN, R., DE VRIES, A. H., STOKER, J., AND VAN VLIET, L. J. Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography. In *MICCAI* (2006), pp. 471–478.
- [68] VEMURI, B., MITICHE, A., AND AGGARWAL, J. Curvature-based representation of objects from range data. *Image and Vision Computing*, 4 (1986), 107–114.
- [69] VINING, D. J. Virtual endoscopy: is it reality? *Radiology* 200 (1996), 30–31.
- [70] VINING, D. J. Virtual colonoscopy. *Gastrointestinal Endosc Clin. N. Am* 7, 2 (1997), 285–291.
- [71] WAN, M., LIANG, Z., KE, Q., HONG, L., BITTER, I., AND KAUFMAN, A. Automatic centerline extraction for virtual colonoscopy. *IEEE Transactions on Image Processing* 21, 12 (2002), 1450–1460.

- [72] WANG, S., WANG, Y., JIN, M., GU, X., AND SAMARAS, D. Conformal geometry and its applications on 3D shape matching, recognition and stitching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 7 (2007), 1209–1220.
- [73] WANG, Y., GUPTA, M., ZHANG, S., WANG, S., GU, X., SAMARAS, D., AND HUANG, P. High resolution tracking of non-rigid 3d motion of densely sampled data using harmonic maps. In *ICCV* (2005), pp. 388–395.
- [74] WANG, Z., LIANG, Z., LI, X., LI, L., LI, B., EREMINA, D., AND LU, H. An improved electronic colon cleansing method for detection of colonic polyps by virtual colonoscopy. *IEEE Transactions on Biomedical Engineering* 53, 8 (2006), 1635–1646.
- [75] WU, C., CLIPP, B., LI, X., FRAHM, J.-M., AND POLLEFEYS, M. 3D model matching with viewpoint-invariant patches (VIP). In *Proc. CVPR* (2008), pp. 1–8.
- [76] YANG, P., AND QIAN, X. Direct computing of surface curvatures for point-set surfaces. In *Proceedings of 2007 IEEE/Eurographics Symposium on Point-based Graphics (PBG)* (2007).
- [77] Z. ZHANG, R. DERICHE, O. F. Q. T. L. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence* 78 (1995), 87–119.
- [78] ZAHARESCU, A., BOYER, E., VARANASI, K., AND HORAUD, R. P. Surface feature detection and description with applications to mesh matching. In *Proc. CVPR* (2009), pp. 373–380.
- [79] ZALIS, M. E., PERUMPILLICHIRA, J., AND HAHN, P. F. Digital subtraction bowel cleansing for CT colonography using morphological and linear filtration methods. *IEEE Transactions on Medical Imaging* 23, 11 (2004), 1335–1343.

- [80] ZHANG, D., AND HEBERT, M. Harmonic maps and their applications in surface matching. In *Proc. CVPR* (1999), pp. 524–530.
- [81] ZOU, G., HUA, J., GU, X., AND MUZIK, O. An approach for intersubject analysis of 3D brain images based on conformal geometry. In *Proceedings of International Conference on Image Processing* (2006), pp. 1193–1196.
- [82] ZOU, G., HUA, J., LAI, Z., GU, X., AND DONG, M. Intrinsic geometric scale space by shape diffusion. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1193–1200.
- [83] ZOU, G., HUA, J., AND MUZIK, O. Non-rigid surface registration using spherical thin-plate splines. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention* (2007), pp. 367–374.

ABSTRACT

MULTI-SCALE AND MULTI-SPECTRAL SHAPE ANALYSIS: FROM 2D TO 3D

by

ZHAOQIANG LAI

December 2011

Advisor: Dr. Jing Hua

Major: Computer Science

Degree: Doctor of Philosophy

Shape analysis is a fundamental aspect of many problems in computer graphics and computer vision, including shape matching, shape registration, object recognition and classification. Since the SIFT achieves excellent matching results in 2D image domain, it inspires us to convert the 3D shape analysis to 2D image analysis using geometric maps. However, the major disadvantage of geometric maps is that it introduces inevitable, large distortions when mapping large, complex and topologically complicated surfaces to a canonical domain. It is demanded for the researchers to construct the scale space directly on the 3D shape.

To address these research issues, in this dissertation, in order to find the multiscale processing for the 3D shape, we start with shape vector image diffusion framework using the geometric mapping. Subsequently, we investigate the shape spectrum field by introducing the implementation and application of Laplacian shape spectrum. In order to construct the scale space on 3D shape directly, we present a novel idea to solve the diffusion equation using the manifold harmonics in the spectral point of view. Not only confined on the mesh, by using the point-based manifold harmonics, we rigorously derive our solution from the diffusion equation which is the essential of the scale space processing on the manifold. Built upon the point-based manifold harmonics transform, we generalize the diffusion function directly on the point

clouds to create the scale space. In virtue of the multiscale structure from the scale space, we can detect the feature points and construct the descriptor based on the local neighborhood. As a result, multiscale shape analysis directly on the 3D shape can be achieved.

AUTOBIOGRAPHICAL STATEMENT

ZHAOQIANG LAI

Zhaoqiang Lai is a Ph.D. candidate in Computer Science at Wayne State University, where he is also a research assistant in the Graphics and Imaging Laboratory. He received his BS degree (2003) in Computer Science and MS degree (2006) in Pattern Recognition and Intelligent System, from Huazhong University of Science and Technology (HUST), Wuhan, China. His research interests include computer graphics, image, visualization, and their applications to computer vision and medical imaging.