

University of Louisville

## ThinkIR: The University of Louisville's Institutional Repository

---

Electronic Theses and Dissertations

---

12-2013

### Geometric modeling of non-rigid 3D shapes : theory and application to object recognition.

Mostafa Abdelrahman  
*University of Louisville*

Follow this and additional works at: <https://ir.library.louisville.edu/etd>



Part of the [Electrical and Computer Engineering Commons](#)

---

#### Recommended Citation

Abdelrahman, Mostafa, "Geometric modeling of non-rigid 3D shapes : theory and application to object recognition." (2013). *Electronic Theses and Dissertations*. Paper 4.  
<https://doi.org/10.18297/etd/4>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact [thinkir@louisville.edu](mailto:thinkir@louisville.edu).

GEOMETRIC MODELING OF NON-RIGID 3D SHAPES: THEORY AND  
APPLICATION TO OBJECT RECOGNITION

By

Mostafa Abdelrahman

B.Sc. 2001, M.Sc. 2006, EE, Assiut University.

A Dissertation

Submitted to the Faculty of the

J. B. Speed School of Engineering of the University of Louisville

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

University of Louisville

Louisville, Kentucky

December 2013

© Copyright 2013 by Mostafa Abdelrahman  
All rights reserved



GEOMETRIC MODELING OF NON-RIGID 3D SHAPES: THEORY AND  
APPLICATION TO OBJECT RECOGNITION

By

Mostafa Abdelrahman  
B.Sc. 2001, M.Sc. 2006, EE, Assiut University.

A Dissertation Approved on

November 20<sup>th</sup> 2013

by the Following Reading and Examination Committee:

---

Aly A. Farag, Ph.D., Dissertation  
Director

---

Thomas L. Starr, Ph.D.

---

John F. Naber, Ph.D.

---

Prasanna K. Sahoo, Ph.D.

---

Tamer Inanc, Ph.D.

---

Moumen A. Elmelegy, Ph.D.

## DEDICATION

I would like to dedicate this dissertation to my mother and adorable wife **Rehab**, who has taking care of our sons *Mohamed, Zyad* and *Omar* alone, while I was working to complete my Ph.D. degree. All I can say is ” *Thank you, my love!*”

## ACKNOWLEDGMENTS

First of all, I would like to thank **Almighty God**, the merciful, the compassionate for all the blessings and for granting me the opportunity to finish and establish this research.

I would like to express my deepest gratitude to my advisor **Professor Aly A. Farag** for his guidance, support, and encouragement during the course of my Ph.D. Both the atmosphere in his group and the wide range of research topics have provided not only an ideal but inspiring environment for my work. I also would like to thank him for giving me the chance to work on interesting projects, meet highly qualified people in the field, and attend several top ranked scientific meetings.

I would like to thank all my dissertation committee members for all the fruitful discussions on this work. I also would like to thank our lab engineer Mr. Mike Miller for being there when things looked bleak, making the lab equipment accessible, and ensuring that the lab runs smoothly. I would like to thank all the members of the computer vision and image processing laboratory (CVIP) at the University of Louisville for their encouragement and support. Special thanks to Dr. Moumen Elmelegy and Dr. Asem Ali for their support.

I would like to thank my parents for their continuous encouragement, and unconditional love. Without them this dissertation would not have been possible. Finally, I would like to thank my adorable wife ,**Rehab**, for her true love, endless encouragement, sharing the difficult moments when things got tough, and for taking care of our sons Mohamed, Zyad, and Omar during the period of my degree.

# ABSTRACT

## GEOMETRIC MODELING OF NON-RIGID 3D SHAPES: THEORY AND APPLICATION TO OBJECT RECOGNITION

Mostafa Aly Ahmed Abdelrahman

November 20<sup>th</sup> 2013

One of the major goals of computer vision is the development of flexible and efficient methods for shape representation. This is true, especially for non-rigid 3D shapes where a great variety of shapes are produced as a result of deformations of a non-rigid object. Modeling these non-rigid shapes is a very challenging problem. Being able to analyze the properties of such shapes and describe their behavior is the key issue in research. Also, considering photometric features can play an important role in many shape analysis applications, such as shape matching and correspondence because it contains rich information about the visual appearance of real objects. This new information (contained in photometric features) and its important applications add another, new dimension to the problem's difficulty.

Two main approaches have been adopted in the literature for shape modeling for the matching and retrieval problem, local and global approaches. Local matching is performed between sparse points or regions of the shape, while the global shape approaches similarity is measured among entire models. These methods have an underlying assumption that shapes are rigidly transformed. And Most descriptors proposed so far are confined to shape, that is, they analyze only geometric and/or topological properties of 3D models.

A shape descriptor or model should be isometry invariant, scale invariant,



be able to capture the fine details of the shape, computationally efficient, and have many other good properties. A shape descriptor or model is needed. This shape descriptor should be: able to deal with the non-rigid shape deformation, able to handle the scale variation problem with less sensitivity to noise, able to match shapes related to the same class even if these shapes have missing parts, and able to encode both the photometric, and geometric information in one descriptor. This dissertation will address the problem of 3D non-rigid shape representation and textured 3D non-rigid shapes based on local features. Two approaches will be proposed for non-rigid shape matching and retrieval based on Heat Kernel (HK), and Scale-Invariant Heat Kernel (SI-HK) and one approach for modeling textured 3D non-rigid shapes based on scale-invariant Weighted Heat Kernel Signature (W-HKS).

For the first approach, the Laplace-Beltrami eigenfunctions is used to detect a small number of critical points on the shape surface. Then a shape descriptor is formed based on the heat kernels at the detected critical points for different scales. Sparse representation is used to reduce the dimensionality of the calculated descriptor. The proposed descriptor is used for classification via the Collaborative Representation-based Classification with a Regularized Least Square (CRC-RLS) algorithm. The experimental results have shown that the proposed descriptor can achieve state-of-the-art results on two benchmark data sets.

For the second approach, an improved method to introduce scale-invariance has been also proposed to avoid noise-sensitive operations in the original transformation method. Then a new 3D shape descriptor is formed based on the histograms of the scale-invariant HK for a number of critical points on the shape at different time scales. A Collaborative Classification (CC) scheme is then employed for object classification. The experimental results have shown that the proposed descriptor can achieve high performance on the two benchmark data sets. An important observation from the experiments is that the proposed approach is more

able to handle data under several distortion scenarios (noise, shot-noise, scale, and under missing parts) than the well-known approaches.

For modeling textured 3D non-rigid shapes, this dissertation introduces, for the first time, a mathematical framework for the diffusion geometry on textured shapes. This dissertation presents an approach for shape matching and retrieval based on a weighted heat kernel signature. It shows how to include photometric information as a weight over the shape manifold, and it also propose a novel formulation for heat diffusion over weighted manifolds. Then this dissertation presents a new discretization method for the weighted heat kernel induced by the linear FEM weights. Finally, the weighted heat kernel signature is used as a shape descriptor. The proposed descriptor encodes both the photometric, and geometric information based on the solution of one equation.

Finally, this dissertation proposes an approach for 3D face recognition based on the front contours of heat propagation over the face surface. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The propagation contours are used to successfully discriminate the various faces. The proposed approach is evaluated on the largest publicly available database of 3D facial images and successfully compared to the state-of-the-art approaches in the literature.

This work can be extended to the problem of dense correspondence between non-rigid shapes. The proposed approaches with the properties of the Laplace-Beltrami eigenfunction can be utilized for 3D mesh segmentation. Another possible application of the proposed approach is the view point selection for 3D objects by selecting the most informative views that collectively provide the most descriptive presentation of the surface.

# TABLE OF CONTENTS

<b>DEDICATION</b>	<b>iii</b>
<b>ACKNOWLEDGMENTS</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>xiii</b>
<b>LIST OF FIGURES</b>	<b>xv</b>
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Shape Representation . . . . .	3
1.1.1 Properties of Shape Descriptor . . . . .	4
1.1.2 Data Representation of Shapes . . . . .	5
1.1.2.1 Point-based Representation . . . . .	5
1.1.2.2 Boundary-based Representation . . . . .	7
1.1.2.3 Volume-based Representation . . . . .	7
1.1.3 Model-based Shape Representation Techniques . . . . .	8
1.1.3.1 Global Feature-based Techniques . . . . .	8
1.1.3.2 Graph-based Techniques . . . . .	9
1.1.3.3 Local Feature-based Techniques . . . . .	9
1.2 Why this Work? . . . . .	10
1.2.1 Motivation Behind this Work . . . . .	12
1.3 Contributions and Proposed Approach . . . . .	13
1.4 Dissertation Organization . . . . .	15

<b>2</b>	<b>EXISTING 2D AND 3D SHAPE DESCRIPTORS</b>	<b>16</b>
2.1	2D Detectors and Descriptors . . . . .	17
2.1.1	Local Features . . . . .	18
2.1.2	Properties of the Ideal Local Feature . . . . .	20
2.1.3	Local Feature Detectors . . . . .	22
2.1.4	Local Feature Descriptors . . . . .	23
2.2	3D Detectors and Descriptors . . . . .	23
2.2.1	3D Feature Detectors . . . . .	26
2.2.1.1	Harris 3D . . . . .	27
2.2.1.2	Mesh DOG . . . . .	27
2.2.1.3	Heat Kernel Feature Detectors . . . . .	27
2.2.2	3D Feature Descriptors . . . . .	28
2.2.2.1	Shape Context . . . . .	28
2.2.2.2	Spin Images . . . . .	28
2.2.2.3	Mesh HOG . . . . .	29
2.2.2.4	Geodesic Distance . . . . .	29
2.2.2.5	ShapeDNA . . . . .	29
2.2.2.6	Heat Kernel Signatures (HKS) . . . . .	30
2.2.2.7	Scale Invariant Heat Kernel Signatures (SI-HKS)	30
2.2.2.8	Scale Invariant Feature Transform for Meshes (mesh- SIFT) . . . . .	31
2.3	Summary . . . . .	31
<b>3</b>	<b>FUNDAMENTALS OF HEAT EQUATION AND THE LAPLACE- BELTRAMI</b>	<b>32</b>
3.1	Fundamentals of Geometry . . . . .	32
3.1.1	Basic Terms in Geometry and Topology . . . . .	33
3.1.2	Length Spaces . . . . .	37
3.1.3	Manifolds . . . . .	38

3.1.4	Embedded Surfaces . . . . .	40
3.1.5	First Fundamental Form . . . . .	41
3.1.6	Curvature and the Second Fundamental Form . . . . .	42
3.1.7	Isometries . . . . .	43
3.2	The Heat Equation . . . . .	45
3.2.1	Initial Boundary Value Problems (IBV) . . . . .	46
3.2.2	The Heat Equation on a Manifold . . . . .	46
3.3	Laplace-Beltrami . . . . .	47
3.3.1	Laplace-Beltrami Operator . . . . .	47
3.3.2	Discrete Laplace-Beltrami Operator . . . . .	48
3.4	Heat Operator and Heat Kernel . . . . .	52
3.4.1	Heat Kernel . . . . .	59
3.4.2	Finite Element Discretization of the Heat Equation . . . . .	60
3.4.3	Properties of the Heat Kernel . . . . .	63
3.4.4	Heat Kernel Signature . . . . .	63
3.4.5	Temperature Distribution Descriptor (TD) . . . . .	64
3.5	Summary . . . . .	64
<b>4</b>	<b>PROPOSED SHAPE DESCRIPTORS</b>	<b>65</b>
4.1	Critical Points Detection . . . . .	67
4.2	Shape Descriptor Based on Heat Kernels . . . . .	67
4.2.1	The Heat Kernel Descriptor . . . . .	70
4.2.2	Sparse Representation of HK . . . . .	72
4.2.3	Experimental Results . . . . .	73
4.2.4	Noise Effect Analysis . . . . .	79
4.3	A Shape Descriptor Based on Scale Invariant Heat Kernels . . . . .	80
4.3.1	Weak And Discrete Formulation of The Heat Equation . . . . .	81
4.3.2	Scale Invariance . . . . .	82
4.3.3	Scale Invariant Heat Kernel Descriptor . . . . .	85

4.3.4	Collaborative Classification . . . . .	85
4.3.5	Experimental Results . . . . .	86
4.3.6	Distorted Data Experiment . . . . .	92
4.3.7	Using Different Classifiers . . . . .	93
4.4	Summary . . . . .	96
<b>5</b>	<b>HEAT DIFFUSION OVER WEIGHTED MANIFOLDS: A NEW DESCRIPTOR FOR TEXTURED 3D NON-RIGID SHAPES</b>	<b>99</b>
5.1	Introduction . . . . .	99
5.1.1	Review of Related Work . . . . .	100
5.1.2	Contribution . . . . .	101
5.2	Heat Kernel on Weighted Manifold . . . . .	102
5.2.1	Finite Element Discretization of the Weighted Heat Equation	103
5.3	Weighted Heat Kernel Signature (W-HKS) . . . . .	107
5.4	Experimental Results . . . . .	108
5.5	Summary . . . . .	110
<b>6</b>	<b>HEAT FRONT PROPAGATION CONTOURS FOR 3D FACE RECOGNITION</b>	<b>113</b>
6.1	Introduction . . . . .	113
6.2	Proposed Approach for 3D Face Recognition . . . . .	116
6.2.1	Facial Landmark Detection . . . . .	117
6.2.2	Heat Front Propagation Contours . . . . .	118
6.2.3	Contours Matching . . . . .	119
6.3	Experimental Results . . . . .	121
6.4	Summary . . . . .	123
<b>7</b>	<b>CONCLUSIONS AND FUTURE WORK</b>	<b>125</b>
	<b>REFERENCES</b>	<b>144</b>

APPENDICES	145
Appendix A: The Derivation of the Heat Equation	145
Appendix B: Analytic Derivation of Equation 3.32	149
Appendix C: Analytic Derivation of Equations 3.34 and 3.34	153
Appendix D: Analytic Derivation of Equation 5.10	155
Appendix E: Analytic Derivation of Equations 5.11, 5.12, and 5.13	159
Appendix F: Iterative Closest Point (ICP) Algorithm	162
NOMENCLATURE	168
CURRICULUM VITAE	169

## LIST OF TABLES

1.1	Requirements of ideal shape representation method. . . . .	4
4.1	Results on the non-rigid world data set. Upper four rows are results using KNN, while lower four rows are results using CRC-RLS. . .	77
4.2	Results on SHREC'11 data set. Upper four rows are results using KNN, while lower four rows are results using CRC-RLS. . . . .	77
4.3	Results on the non-rigid world data set using the proposed descriptors computed at only one critical point. Upper row using KNN, while lower row using CRC-RLS. . . . .	77
4.4	Results on SHREC'11 data set with respect to size (N) of the SRCP-TD descriptor using the CRC-RLS. . . . .	78
4.5	Results on non-rigid world data set with different NSR of the SRCP-TD descriptor using the CRC-RLS. . . . .	80
4.6	Results on non-rigid world data set. The proposed method is compared to TD [1], Shape-DNA [2], and Shape-Google [3]. . . . .	88
4.7	Results on SHREC'11 data set. The proposed method is compared to TD [1], Shape-DNA [2], and Shape-Google [3]. . . . .	88
4.8	Performance versus white noise, shot-noise, and scale in three severity levels of the proposed CP-HK descriptor using the RCR classifier compared to those of Shape Google [3] and Shape-DNA [2] (1.00 mean 100%). . . . .	92



4.9	Average accuracy measures on the non-rigid world data set with different classifiers. For all measures, the higher the better except the FP rate, where the lower the better. Upper four rows are the proposed approach, middle are for the TD method, while the lower four rows are for the Shape-DNA. . . . .	95
4.10	Average accuracy measures on the SHREC'11 data set with different classifiers. For all measures, the higher the better except the FP rate, where the lower the better. Upper four rows are the proposed approach, middle are for the TD method, while the lower four rows are for the Shape-DNA. . . . .	96
4.11	Average accuracy measures on the SHREC'11 data set with different noise levels, Upper is the proposed method, Lower is Shape-DNA.	96
4.12	Average accuracy measures on the SHREC'11 data set with different shot-noise levels. . . . .	97
4.13	Average accuracy measures on the SHREC'11 data set with different scale levels. Upper is the proposed method, Lower is Shape-DNA.	97
5.1	Results on SHREC'13 data set compares the proposed descriptor (W-HKS1) against the (A2) based on Scale Invariant Heat Kernels combined with the color histogram. . . . .	110
5.2	Results on SHREC'13 data set lists the average measures on SHREC'13 data set. The table compares the proposed descriptor (W-HKS2) against five different methods. . . . .	111

## LIST OF FIGURES

1.1	(a) Point-based shape representation using cloud of points, (b) Surface-based shape representation using polygonal mesh, (c) Volume based shape representation using grid of voxels. . . . .	6
1.2	(a) Microsoft Kinect, (b) Example of 3D reconstruction generated by the Kinect. . . . .	7
2.1	Examples of 2D object retrieval (top), 3D shape retrieval (middle), and textured 3D shape retrieval (bottom). Shown on the left is a query, and on the right, examples of retrieved or matched objects. Transformations shown in image retrieval are viewpoint variation, different color, and background variation, in shape retrieval different non-rigid shape deformations are shown, and for textured 3D shape retrieval similar shapes with similar texture are shown. . . .	17
2.2	Example of a 3D shape to the left and the detected points to the right using Harris 3D detector. . . . .	26
2.3	Examples of detected features using Mesh DOG on individual 3D human teeth. . . . .	28
2.4	Example of feature matching using meshSIFT on 3D colon parts.	30
3.1	Point $\mathbf{x}$ is an interior point of $A$ , point $\mathbf{y}$ is an exterior point of $A$ , while $\mathbf{z}$ is a boundary point of $A$ . . . . .	35
3.2	Two types of metrics. . . . .	37
3.3	The two types of manifolds. . . . .	39
3.4	3D surface parametrization, tangent plane, and normal. . . . .	40

3.5	Illustration of first fundamental form. . . . .	41
3.6	illustration of maximum principal curvature $\kappa_1$ and the minimum principal curvature $\kappa_2$ in the upper row, and Gaussian curvature $K$ and mean curvature $H$ in the lower row. . . . .	44
3.7	Two types of isometries, the original shape $(\mathbf{X}, \alpha)$ , the Euclidean isometric copy $(\tau\mathbf{X}, \alpha)$ , and the deformed isometric copy $(\mathbf{X}, \beta)$ . . . . .	45
3.8	Discretization of the Laplace-Betrami operator using cotangent weights. Left: a vertex $\mathbf{x}_i$ and its adjacent faces Middle: the definition of the angles $\alpha_{ij}$ and $\beta_{ij}$ Right: the definition of the area $A_i$ . . . . .	51
3.9	The first non-zero eigenfunctions of the Laplace-Beltrami operator of cat shape from the nonrigid world data set [4, 5] with two different views. Colors represent the values of the eigenfunction at each point of the shape. This illustration shows that the first non-trivial eigenfunctions are invariant to the shape deformation. . . . .	53
3.10	The first non-zero eigenfunctions of the Laplace-Beltrami operator of lion shape with two different views. Colors represent the values of the eigenfunction at each point of the shape. This illustration shows that the first non-trivial eigenfunctions are invariant to the shape deformation. . . . .	54
3.11	The first four eigenfunctions of the Laplace-Beltrami operator of human shapes with different deformation. This illustration shows that the eigenfunctions are invariant to the shape deformation and can be used for shape segmentation as [6]. . . . .	55
3.12	The first three eigenfunctions of the Laplace-Beltrami operator of women shapes with different deformation. This illustration shows that the eigenfunctions are invariant to the shape deformation and can be used for shape segmentation as [6]. . . . .	56

3.13	Spectral projection onto the first 3 eigenfunctions of the Laplace-Beltrami of human shape with different deformation. Left: the 3D shape, Middle: the spectral projection as a surface, Right: the spectral projection as a triangulated mesh. . . . .	57
3.14	Spectral projection onto the first 3 eigenfunctions of the Laplace-Beltrami of dog shape with different deformation. Notice the difference in the projection over this of the human shape in Figure 3.13 Left: the 3D shape, Middle: the spectral projection as a surface, Right: the spectral projection as a triangulated mesh. . . . .	58
4.1	The three parts of object matching algorithms. . . . .	65
4.2	Critical points detected for different shapes with different deformation. Number below each shape represents the total number of shape critical points. Colors indicate the average temperature induced from these critical points throughout all shape vertices. . . . .	68
4.3	HKS induced from one critical point detected at the nose of the dog shape at different time samples. The shown numbers indicate the time samples $t$ . Colors represent the values of the heat induced from the source point to each vertex of the shape. For a smaller amount of time, the signature is more local. As time progresses, the signature becomes more global. . . . .	71
4.4	Construction of the proposed descriptor ( $N$ is the number of object vertices). . . . .	72
4.5	Some shape retrieval results on the nonrigid world data set. Left: Queries. Right: First 10 matches using the SRCP-TD descriptor. The color represents the mean temperature induced from only the detected critical points. The detected sparse points are shown in yellow. The green-colored shapes are wrong matches. . . . .	75

4.6	Shape retrieval results of SHREC 2011 data set. Left: queries. Middle: First 15 matches using the SRCP-TD descriptor. Right: matches from 30 to 35. The color represents the mean temperature induced from only the detected critical points shown in yellow. . . . .	76
4.7	Single critical points detected for several shapes of the non-rigid world data set. . . . .	78
4.8	noise-corrupted models. from left to right 0.0, levels 1, 2, and 3 <i>NSR</i> . . . . .	79
4.9	Construction of the proposed descriptor ( $N$ is the number of object vertices). . . . .	81
4.10	Construction of the scale-invariant heat kernel under several noise levels. (Left) heat kernel computed at a point on a shape at different time (red) and the scaled heat kernel (in green) computed at a corresponding point on a shape scaled up by a factor of 3. (Middle) The amplitude of the first 15 fourier transform components of $ H(w) $ for the two heat kernels (again in red and green) using the proposed method showing complete overlapping specially in the first four noise levels. (Right) Similar but using the method in [3]. First row shows signal without noise. Then noise level is increased in the subsequent rows. . . . .	84
4.11	Some shape retrieval results on the nonrigid world data set. Left: Queries. Right: First 10 matches using the CP-HK descriptor. The detected sparse points are shown in yellow. The light green-colored shapes are wrong matches. . . . .	89
4.12	Shape retrieval results of SHREC'11 data set. Left: queries (colored). Then First 20 matches using the proposed descriptor. The light green colored shapes represent wrong matches. The detected critical points are shown in yellow. . . . .	90

4.13	An illustration of confusion matrix ( $600 \times 600$ ) for the SHREC'11 data set. Each row in this matrix represents the first 20 match retrieval results for each shape in the data set(600 shapes). Note that the shapes belonging to the same class are sorted to be next to each other. . . . .	91
4.14	Some shape retrieval results for different shapes with different noise levels. Left: queries. Right: the first 5 matched shapes. . . . .	93
4.15	Some shape retrieval results for different shapes with different shot noise levels. Left: queries. Right: the first 5 matched shapes. . . .	94
4.16	Some shape retrieval results for different shapes with different missing parts. Left: queries. Right: the first 5 matched shapes. . . . .	94
5.1	Construction of the proposed descriptor. . . . .	102
5.2	Sample of 4 different classes of the SHREC'13 benchmark [7] with different texture and deformation show the challenge of the data set. For each null shape there are a transformed versions of it that include non-rigid deformation, nonmetric-preserving deformations and additive Gaussian noise perturbation. . . . .	109
5.3	Shape retrieval results of SHREC'13 data set. Left: queries. Right: First 15 matches using the CHKS descriptor. . . . .	111
5.4	Shape retrieval results of SHREC'13 data set. Precision-recall curves for all methods. . . . .	112
6.1	The basic idea of the proposed approach: The heat propagation front contours for a virtual heat source at some facial landmarks (nose tip here) are used as discriminating features between different faces. First and second faces are for the same subject thus having similar contours, while the third face has different contours because it is for another subject. . . . .	114

6.2	Proposed approach: Given a depth image, triangulation to get 3D mesh, feature point extraction, heat kernel reconstruction for each feature point, 3D contours extraction, then contour sampling. . .	117
6.3	Facial landmarks detected on sample faces from Texas 3D faces data set. . . . .	118
6.4	The first four non-zero eigenfunctions of the Laplace-Beltrami operator of human faces . . . . .	119
6.5	HK induced from three different landmarks point detected on the face shape at seven different time samples. Colors represent the values of the heat induced from the source point to each vertex of the shape. For small time, the HK is more local. As time progresses, the HK becomes more global. . . . .	120
6.6	Upper row: The selected 12 landmarks, Lower row: The reconstructed one contour around each point. These 12 contours are used as a descriptor for the 3D face. . . . .	121
6.7	Samples of Texas 3D faces data set [8], from top to bottom The RGB image, the depth image and the 3D mesh. . . . .	122
6.8	3D face recognition framework. . . . .	123
6.9	The CMC curves, for the eigensurfaces, fishersurfaces, ICP algorithms, anthroface 3D algorithm based on 25 manually located points, the anthroface 3D algorithm that employed 10 automatically located points [8], and the proposed approach HFPC. . . .	124
A-1	Thin bar of a homogenous material, any position along the bar is denoted as $x$ , and the length of the bar is denoted as $L$ such that $0 \leq x \leq L$ . . . . .	145

B-1	Discrete triangulated surface. Left:A vertex $\mathbf{x}_i$ and its adjacent faces Middle:the definition of the angles $\alpha_{ij}$ and $\beta_{ij}$ Right: the definition of the interior angles $\alpha_i, \alpha_j, \alpha_k$ of triangle $T$ at vertices $i, j, k$ respectively. . . . .	149
B-2	Shearing a regular triangle to a right angle triangle by changing variables . . . . .	150
D-1	Shearing a regular triangle to a right angle triangle by changing variables. . . . .	155
E-1	The definition of the angles $\alpha_{ij}$ and $\beta_{ij}$ and the weights $h_{T_\alpha}$ and $h_{T_\beta}$ .	160



# CHAPTER 1

## INTRODUCTION

In everyday experience, people are often unaware of how complex the shape analysis performed by their brain is. Recognizing objects and making a decision requires the analysis of the geometric as well as the photometric information of these objects. The human visual system represents and easily recognizes deformable three-dimensional objects. Various psychophysical and physiological studies have indicated that the human visual system encodes three dimensional objects as multiple viewpoint-specific two-dimensional representations with appropriate depth information. Human visual performances are still superior to that of computer vision in many aspects.

For computers, computer vision attempts to imitate the ability of human vision, and aims to make the computer see as a human. A lot of research has been done in the area of computer object recognition (or classification), to find an object-invariant representation, and to search for a robust feature detector and an invariant model for 3D objects, especially non-rigid objects.

There are a great variety of shapes produced as a result of deformations of a non-rigid object. Modeling these non-rigid shapes is a very challenging problem in the area of computer vision. Being able to analyze the properties of such shapes and describe their behavior is the key issue in research.

Geometric modeling is fundamental for visual computing because it provides shape representation for geometric objects by trying to understand their

shapes properties. Geometrical representation of shapes take many forms: two-dimensional, three-dimensional point clouds, triangular meshes, and parametric or implicit surfaces.

After modeling each shape it is needed to look for a distance measure between these shape models. If this distance is small, this conclude that the shapes are similar. In case of non-rigid shapes finding a discriminant distance measure is critical due to the fact that an object can take many forms as a result of its deformations, where the same object shapes seem to be dissimilar due to large deformation.

To be able to correctly check the similarity of non-rigid shapes, it is needed to look for the properties that distinguish between these shapes. Such properties are called deformation invariant signatures, and the similarity measure based on these properties is called deformation-invariant similarity. So deformation invariant signatures and deformation-invariant similarity measure is needed for non-rigid 3D objects.

Recently, many sensors are able to acquire the color information besides the 3D shape. Also multiple-view stereo techniques are able to recover both geometric and photometric information. These photometric features can play an important role in many shape analysis applications, such as shape matching and correspondence because they contain rich information about the visual appearance of real objects. This new requirement and its important applications adds another, new dimension to the difficulty of the problem. Most descriptors proposed so far are confined to shape, that is, they analyze only geometric and/or topological properties of 3D models. Therefore more effort is needed to consider color in addition to shape in object representation and description.

Non-rigid objects are all around, and there is no particular problems in dealing with them in the daily lives. The world is full of objects that, due to their physical properties, are non-rigid and therefore can be deformed and bent.

Non-rigid shapes appear at all scales in nature, for example, the human body, its organs and tissues. However the mathematical tools available for the description and analysis of these deformable objects are few, and only relatively recently researchers in graphics and computer vision have started paying due attention to them. During the past decades, a large number of shape matching and searching techniques have been developed (e.g. using moments, spherical harmonics, or Reeb graphs,...,etc.). Evaluation of 3D shape modeling methods, with respect to other requirements, can be found in [9, 10, 11, 12, 13].

## 1.1 Shape Representation

Recently, using 3D object's data has become more important in the area of computer vision, as recognition systems based on 3D models are less sensitive, or may be invariant, to lighting conditions and pose variations as compared to 2D models. The emergence of laser/lidar sensors, reliable multi-view stereo techniques and more recently consumer depth cameras have made the acquisition of 3D models easier than before. Also recent advances in 3D hardware acquisition devices, such as 3D cameras, laser scanners, magnetic resonance (MR), and computed tomography (CT), made it easy to reconstruct high-quality 3D models, which have to be analyzed or visualized in a concise way. While 3D models can be described by their color, texture, and shape information, both color and texture (photometric data) are not always available, depending on the reconstruction method or acquisition technique. Therefore, shape is the lowest common denominator in describing 3D objects.

Today understanding shape is of great importance in pattern analysis and machine intelligence. One of the major goals of computer vision and machine intelligence is the development of flexible and efficient methods for shape representation or the creation of a shape descriptor or signature for shape matching. The descriptor captures the properties of the shape that distinguish it from shapes

<b>Requirement</b>	<b>Description</b>
<i>Isometry invariant</i>	Isometric shapes should have the same descriptor regarding the objects given representation and location.
<i>Scale invariant</i>	For some applications, it is necessary that the descriptor is independent of the objects size, therefore the descriptor should optionally be scale invariant.
<i>Similarity</i>	Similarly shaped objects should have similar descriptors.
<i>Completeness</i>	<i>(or shape-awareness)</i> Descriptor should give a complete characterization of the shape, thus representing the shape uniquely.
<i>Scope</i>	It must be able to describe all classes of shapes.
<i>Hierarchical</i>	It must describe the shape at multiple scales in a similar fashion to a hierarchical structure (i.e, provides a coarse to fine level of detail).
<i>Accessibility</i>	It must be easy to implement.
<i>Efficiency</i>	It must be computationally efficient, or the time and space needed to compute those descriptors should be reasonable.

Table 1.1: Requirements of ideal shape representation method.

belonging to other classes.

Many geometric shape invariants (e.g. circumference, surface area, volume, bounding sphere or eigenvalues of the inertia tensor) have strong limitations, and are not invariant to shape deformation.

### 1.1.1 Properties of Shape Descriptor

Evaluating the quality of a shape representation is not a trivial task because the method of choice depends on the properties of the described shape and the used application. Shape descriptor should have as many of the properties in Table 1.1 as possible [2], [14].

M. Reuter [2] Defined the properties Isometry and Congruency as: Two geometric objects are isometric if a homeomorphism from one to the other exists preserving (geodesic) distances, i.e. mapping curves to curves with equal arc length. This homeomorphism is then called an isometry.

This is right if the geodesic distances is considered as a deformation invari-

ant representation. But this work considers deformation (or transformation) that does not preserve the length of the curves. So Isometry in general can be defined as:

**Definition 1 *Isometry:*** *Two geometric objects are isometric if there exists a distance-preserving map between metric spaces.*

A more precise definition will be given in Section 3.1.7 then and Congruency can be defined as:

**Definition 2 *Congruency:*** *Two geometric objects are congruent if they can be transformed into each other by rigid motions (translations and rotations) as well as reflections.*

## 1.1.2 Data Representation of Shapes

The 3D shapes can be represented in different levels: The first level is a set of points in 3D space. The second level is the shape boundary. And the third level is the volume of the shape.

### 1.1.2.1 Point-based Representation

One important way, among other possibilities, to represent the acquired 3D model is a 3D point cloud, which can be developed into a surface through meshing. The boundary of the object is described by either a cloud of points as shown in Figure 1.1 (a) or range images. Range images are similar to intensity images in the sense that they capture the shape from one point of view except that the color information of the pixel carries the depth information of the surface point from the camera.

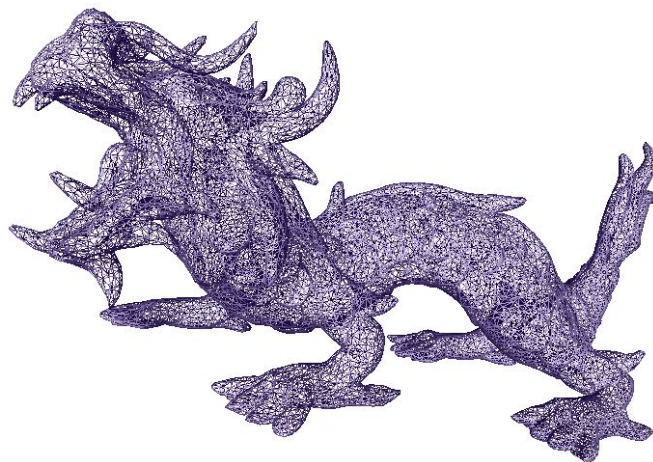
In general, point-based representation lacks the structure information of a shape but it is enough for visualization purposes.



(a)



(b)



(c)

Figure 1.1: (a) Point-based shape representation using cloud of points, (b) Surface-based shape representation using polygonal mesh, (c) Volume based shape representation using grid of voxels.

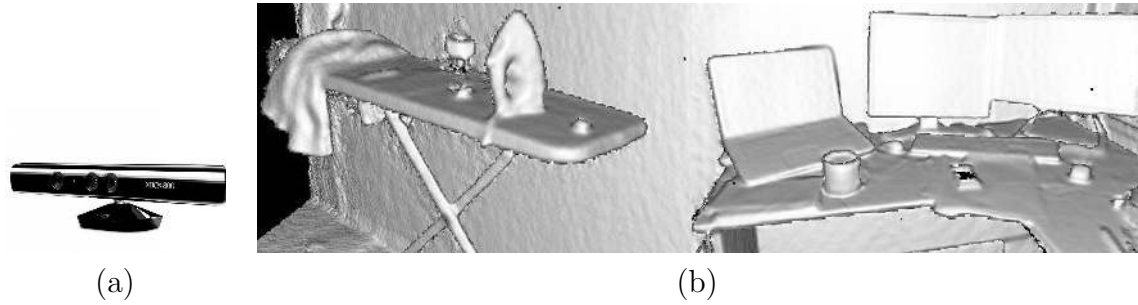


Figure 1.2: (a) Microsoft Kinect, (b) Example of 3D reconstruction generated by the Kinect.

Kinect (see Figure 1.2 (a)) is an amazing sensor that can be used to generate range and RGB images. The depth data from Kinect can be used to track the 3D pose of the sensor and reconstruct, geometrically precise, 3D models of the physical scene in real-time, see Figure 1.2 (b). Also it can provide the photometric information in the RGB image.

#### 1.1.2.2 Boundary-based Representation

A 3D object can be represented in terms of its boundary or surface, which is the common representation in a computer aided design (CAD) and in computer graphics industries. The boundary data can be described by polygonal meshes as shown in Figure 1.1 (b), parametric forms, or implicit surfaces.

#### 1.1.2.3 Volume-based Representation

A 3D shape is represented by the volume it occupies. The volumetric data is described by voxels as shown in Figure 1.1 (c) or solid primitives. A voxel is the minimum 3D unit of a volume. This representation is commonly used in the field of computer vision and medical imaging due to the nature of the acquisition process.

### 1.1.3 Model-based Shape Representation Techniques

Although literature is rich with several shape descriptors definitions, they can be classified under two main categories: model-based or view-based. Model-based methods directly make use of the 3D data, while view-based methods store a number of 2D projections of the 3D object. Previous investigations, such as [15], [16] demonstrate that view-based methods with pose normalization pre-processing can achieve better performance in retrieving rigid shapes, but at higher storage and computational cost than many other 3D descriptors.

Model-based techniques can be categorized into global feature-based techniques, graph-based techniques, and local feature techniques.

**Definition 3 *Descriptor:*** *A feature descriptor is a representation of a feature, or a region around a feature, computed from attributes such as intensity, color or texture of the feature or the region.*

**Definition 4 *Global feature:*** *Global features describe objects as a whole using a single descriptor.*

#### 1.1.3.1 Global Feature-based Techniques

Global feature-based methods use global properties of the object such as moments and spherical harmonics to represent its shape. Moments were used in mechanics for purposes other than shape description.

The advantage of moment based-methods is that they are mathematically concise. The disadvantage is that it is difficult to correlate high-order moments with shape features. Spherical harmonics [17] is another alternative, in which a linear combination of the spherical harmonics are used to describe the surface of the object. Spherical harmonics are a decomposition of a spherical function by finding the Fourier transform on a sphere.



The spherical harmonic coefficients can be used to reconstruct an approximation of the original object at different levels.

In general, global feature-based methods fail to capture the local details of a shape and are not very robust [18].

### 1.1.3.2 Graph-based Techniques

Since the topology of an object is an important shape characteristic, graph-based methods aim at reducing the shape into a graph that captures its topology. The major advantage in representing a 3D model as a topological graph is that it allows representation at multiple levels of detail and hence facilitates analysis and matching of local geometry.

**B-Rep Graph:** Boundary representation (B-Rep) is a common representation in CAD and computer graphics industries, where the object's boundary is represented as a graph. The nodes of the graph represent the set of bounding surfaces while the edges represent the intersecting curves between corresponding surfaces.

**Skeletal Graph:** Skeletal graph-based techniques compute a 1D skeleton of an object and then convert it into a skeletal graph as its shape representation. A skeleton is a compact representation of a shape that maintains its topology. Each point of the skeleton is associated with a radius function, that is the radius of its maximal disk/sphere. Therefore, the shape can be represented with less information than the original one.

Generally, graph-based techniques suffer from missing parts and noise artifacts.

### 1.1.3.3 Local Feature-based Techniques

Although global shape descriptors have shown good performance on many data sets, they have an underlying assumption that shapes are rigidly transformed.

Different approaches used local feature detection and a local descriptor to describe 3D shapes, such as spin images [19], surface point signature [20], contour and edge detection features [21], local patches [22], and conformal factor [23]. However these methods cannot deal with the non-rigid shape deformation, and cannot cover the properties of the desired shape descriptor.

The problem of non-rigid shapes deformation needs more work to compensate for the degrees of freedom resulting from local deformations. Early work by Elad and Kimmel [4, 24] proposed modeling shapes as metric spaces with the geodesic distances as an intrinsic metric, which are invariant to inelastic deformations. Ling and Jacobs [25] and Bronstein et al. [26] used this framework with a metric defined by internal distances in 2D shapes. Reuter et al. [2, 27] used the Laplacian spectra as intrinsic shape descriptors, and they employed the Laplace-Beltrami spectra as 'shape-DNA' or a numerical fingerprint of any 2D or 3D manifold (surface or solid). They proved that 'shape-DNA' is an isometry-invariant shape descriptor. Rustamov [28] used an isometry-invariant shape representation in the Euclidean space, and then histograms of Euclidean distances to compare between shapes.

## 1.2 Why this Work?

One of the most challenging settings is the case of non-rigid or deformable shapes, in which the class of transformations may be very wide due to the capability of such shapes to bend and assume different forms. When trying to adapt feature-based approaches applied to 2D images to 3D shapes, the following have to be considered:

- First, the invariance type in non-rigid 3D shapes is different than these in images. In images typical feature detectors and descriptors should be invariant to affine transformations. In case of non-rigid shapes, the varia-

tion of transformations is much larger, including changes in pose, bending, and connectivity changes. Since many natural shape deformations such as bending can be approximated by isometries, so the shape descriptors depend on intrinsic properties of the shape. This makes it invariant to such deformations.

- Second, shapes are not rich in features like 2D images, making it harder to detect a large number of repeatable and stable key-points. This difficulty motivates a lot of researchers to avoid feature detection at all and use dense descriptors instead.
- Finally, shapes are often represented as triangular meshes, point clouds, voxels, level sets, etc. Therefore, local features is desirable to be computable across multiple representations.

As such, the area of shape modelling and representation still needs more work, and this is the domain of the current dissertation. Also, considering photometric features can play an important role in many shape analysis applications, such as shape matching and correspondence because it contains rich information about the visual appearance of real objects. This new information (contained in photometric features) and its important applications adds another, new dimension to the problem difficulty. Most descriptors proposed so far are confined to shape, that is, they analyze only geometric and/or topological properties of 3D models. Therefore more efforts need to be done to consider color in addition to shape in object representation and description. This dissertation will consider also the representation of textured shapes in order to develop an efficient descriptor that combines the color information as well as the geometric shape information. The sought representation should cope with non-rigid transformations, which is a key requirement for many target applications.

### 1.2.1 Motivation Behind this Work

Although global shape descriptors have shown good performance on many data sets, they have an underlying assumption that shapes are rigidly transformed. So global methods cannot deal with the non-rigid shape deformation. Other methods use local feature detectors and local descriptors to describe 3D shapes, but these methods cannot deal with the non-rigid shape deformation, cannot cover the properties of the desired shape descriptor, and suffer a lot when shapes are corrupted with noise or in the case of partial shape matching. Recent work by Bronstein et al [3] uses the Heat Kernel Signature (HKS) and scale-invariant heat kernel signature as a shape descriptor. However their method to handle the scale variation is sensitive to noise, and their descriptor does not perform well for partial shape matching or in case of shapes with missing parts.

A shape descriptor or model that has as many of the properties in Table 1.1 as possible is needed. This descriptor should be able to deal with the non-rigid shape deformation, should handle the scale variation problem with less sensitivity to noise, should be able to match shapes related to the same class even if these shapes have missing parts.

Recently, taking the photometric information into account to calculate a 3D shape descriptor has attracted more research. The work by Kovnatsky et al. [29] is the most related work to this work; it uses the diffusion geometry framework for the fusion of geometric and photometric information in local and global shape descriptors. Their construction is based on an *ad hoc* definition of a diffusion process on the shape manifold embedded into a high-dimensional space where the embedding coordinates represent the photometric information. Their method fails to provide a mathematical justification for their proposed heat kernel framework or the proposed discretization method. So a shape descriptor or model that encodes both the photometric, and geometric information is needed.

### 1.3 Contributions and Proposed Approach

The construction of a feature-based representation of an image typically consists of two stages, feature detection and feature description, often combined into a single algorithm. The main goal of a feature detector is to find stable points or regions in an image that carry significant information on one hand and can be repeatedly found in transformed versions of the image on the other. The same methodology will be used with the 3D shapes.

For the feature detection stage, and as a first contribution of this dissertation, a new method for feature detection on 3D shapes is proposed. A small number of critical points are detected on the shape surface. An observation shows that shapes belonging to the same class consistently have almost the same number of critical points, whereas these numbers differ from one class to another. As such, this number can be used as one of the discriminatory features between the different classes.

The next stage is feature description. A feature descriptor uses a representation of local image information in the neighborhood of each feature point. As a second contribution, this dissertation proposes two new feature descriptors, one based on the Heat Kernels (HK) [30], and another constructed from scale-invariant heat kernels [31]. The heat kernels are calculated at the detected critical points for different scales. This reduces the size or dimensionality of the descriptors, yet does not sacrifice the accuracy of the presentation. Evaluate the proposed descriptors will be given on the problem of 3D shape retrieval and object classification. Different classification methods are assessed for that goal.

This dissertation also introduces, for the first time, a mathematical framework for the diffusion geometry on textured shapes. It presents an approach for shape matching and retrieval based on a weighted heat kernel signature. This dissertation shows how to include photometric information as a weight over the shape

manifold, also proposes a novel formulation for heat diffusion over weighted manifolds. Then it presents a new discretization method for the weighted heat kernel induced by the linear Finite Elements Method (FEM) weights. This dissertation also proposes a new method to introduce the scale invariance for the weighted heat kernel signature. Finally, the weighted heat kernel signature is used as a shape descriptor. The proposed descriptor encodes both the photometric, and geometric information based on the solution of one equation.

Finally, this dissertation proposes an approach for 3D face recognition, in the framework of the Biometric Optical Surveillance System (BOSS) in the Computer Vision and Image Processing Laboratory (CVIP Lab.), based on the front contours of heat propagation over the face surface. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The propagation contours are used to successfully discriminate the various faces. The proposed approach is evaluated on the largest publicly available database of 3D facial images and successfully compared to the state-of-the-art approaches in the literature.

As such, the specific objectives of this dissertation can be summarized as follows:

- To review literature for the existing methods for non-rigid shape modeling.
- To study the solution of the heat equation and the reconstruction of the heat kernels as a promising direction for shape representation.
- To design a feature detector for non-rigid 3D shapes.
- To design a shape descriptor based on the heat kernel and scale-invariant heat kernels.
- To design a framework for the diffusion geometry on textured shapes.

- To study the application of using the front propagation contours for 3D face recognition.
- To evaluate the proposed methods on existing benchmark data sets.

## 1.4 Dissertation Organization

This document is organized as follows: **Chapter 2** reviews the existing 2D and 3D descriptors. **Chapter 3** gives an introduction to the geometry fundamentals, and discusses the definition of basic terms in geometry and topology, the length spaces, isometries, manifolds, embedded surfaces, first fundamental form, curvature and the second fundamental form. It also details the derivation of the heat equation. Then it describes the Initial Boundary Value Problem (IBV). Also it gives the details of the the Laplace-Beltrami operator, its properties, the discretization of the Laplace-Beltrami operator, and how to use it to solve the heat equation IBV problem on a manifold. After that it talks about the solution of the heat equation and the basics of diffusion on Riemannian manifolds that are necessary to define the heat kernel, and how to solve for the heat kernels using Finite Element Method (FEM) approximation. **Chapter 4** introduces two proposed approaches for shape matching and retrieval based on heat kernel (HK), and scale-invariant heat kernel (HK). **Chapter 5** presents in detail the heat diffusion over weighted manifolds as a new descriptor for textured 3D non-rigid shapes. **Chapter 6** presents a framework for 3D face recognition based on the front propagation contours of heat kernels. **Chapter 7** concludes the proposal and gives insight for extensions and future work to be tackled during this research.

## CHAPTER 2

# EXISTING 2D AND 3D SHAPE DESCRIPTORS

The fast development of 3D data rendering and acquisition techniques brings high attention to the applications related to 3D models (security, multimedia, and biometrics). This development attracted both the scientific and the engineering communities to the problem of shape modeling. The modeling and retrieval of 3D shapes remains a challenging problem. Now, the most important objective is to be able to simply represent 3D objects, and efficiently carry out operations such as object classification, recognition and retrieval.

An active area of research in the image domain is image retrieval: the problem of finding images depicting similar scenes or objects. Images have significant variability similar to three-dimensional shapes, Figure 2.1, and the aim of a successful retrieval approach is to be invariant to such changes, while maintaining high discriminative power. Significant advances have been made in designing efficient image retrieval techniques (see an overview in [32]). Many approaches were recently proposed to describe a 3D object. Some of these approaches are extended from the 2D image application to the 3D domain. Other methods are mainly built for 3D shapes [33]. A 3D shape retrieval algorithm should be able to retrieve shapes belonging to the same class even with different deformations. When considering the photometric information the 3D shape retrieval algorithm should be able to retrieve shapes with similar texture first as shown in Figure 2.1. This chapter will review the existing 2D and 3D descriptors especially local feature



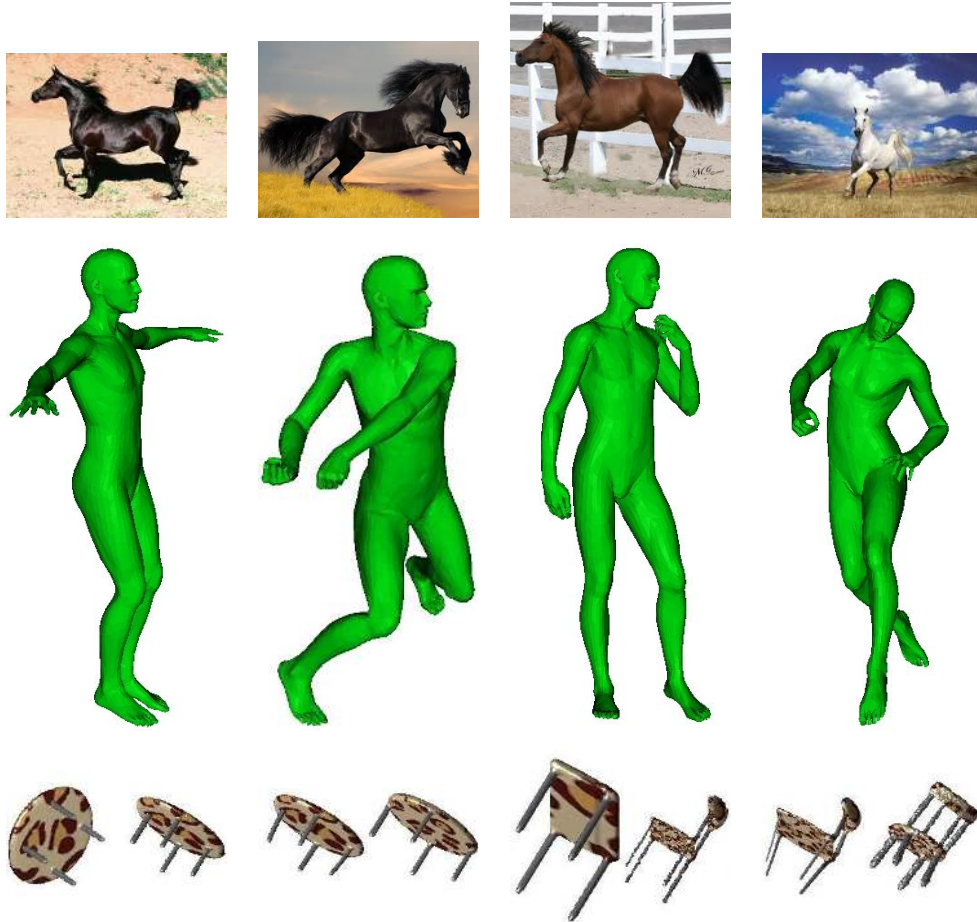


Figure 2.1: Examples of 2D object retrieval (top), 3D shape retrieval (middle), and textured 3D shape retrieval (bottom). Shown on the left is a query, and on the right, examples of retrieved or matched objects. Transformations shown in image retrieval are viewpoint variation, different color, and background variation, in shape retrieval different non-rigid shape deformations are shown, and for textured 3D shape retrieval similar shapes with similar texture are shown.

which is the domain of the current dissertation.

## 2.1 2D Detectors and Descriptors

The state-of-the-art image matching algorithm consists of three parts: feature detector, feature descriptor, and feature matching. In some cases, using one feature that has a strong discriminatory is sufficient to distinguish between objects. For example, the color of the objects can be used as a feature to dis-

tinguish between objects. However, in many cases one feature is not enough for recognition. For example, fruits can be distinguished using the shape and color, and color is not a distinctive feature to distinguish between two flowers. In this case, color and texture may be combined together to provide this sort of distinction. Among the methods used for object description are: Local invariant features, global features, image segments, and exhaustive and random sampling of features.

### 2.1.1 Local Features

Local (invariant) features are a powerful tool, that have been applied successfully in a wide range of systems and applications. The features must be at least partially invariant to illumination, 3D projective transforms, and common object variations. On the other hand, the features must also be sufficiently distinctive to identify specific objects among many alternatives. The difficulty of the object recognition problem is due to the lack of success in finding such image features. However, recent research on the use of dense local features e.g., Schmid and Mohr [34] has shown that efficient recognition can often be achieved by using local image descriptors. Some commercial object recognition systems depend on correlation-based template matching. When object pose and illumination are not controlled enough, then template matching becomes computationally infeasible. When there are object rotation, scale, illumination, and 3D pose, it will be even more challenging.

This problem become more challenging when dealing with partial visibility and large model databases. Instead of searching all image locations for matches the idea is to extract features from the image that are partially invariant to the image formation then process and match those features. Many candidate feature types have been proposed, including line segments [35], groupings of edges [36], and regions [37], among many other works. While these features have worked well

for certain object classes, they are often not detected frequently enough or with sufficient stability to form a basis for reliable recognition. There has been recent work on developing more dense image features. One approach has been to use a corner detector around which local image properties can be measured.

Zhang et al. [38] used the Harris corner detector to identify feature locations for epipolar alignment of images taken from differing viewpoints. Rather than attempting to correlate regions from one image against all possible regions in a second image, large savings in computation time were achieved by only matching regions centered at corner points in each image. For the object recognition problem, Schmid and Mohr [34] also used the Harris corner detector to identify interest points, and then created a local image descriptor at each interest point from an orientation-invariant vector of Derivative-of-Gaussian image measurements. These image descriptors were used for robust object recognition. The corner detectors used in these previous approaches have a major failing, which is that they examine an image at only a single scale. As the change in scale becomes significant, these detectors respond to different image points. Also, since the detector does not provide an indication of the object scale, it is necessary to create image descriptors and attempt matching at a large number of scales.

Some current successful methods for object classification learn and apply quite precise geometric constraints on feature locations [39], others ignore geometry and use a "Bag-of-Features (BoF)" approach that ignores the locations of individual features [39], [40]. Intermediate approaches retain some coarsely-coded location information [41] or record the locations of features relative to the object center [42], [43].

Large database image recognition refers to the task of correctly matching a query image to an image of the same object selected from a large database. In this context large refers to image sets where the amount of data exceeds what can be stored in available memory. Conventional approaches which store individual

local descriptors for each image [36] are no longer suitable as the number of images rises into the millions or higher. One solution to this problem was proposed by Sivic and Zisserman [44], in which image descriptors are quantized into visual words. Quantized matching is performed using a Bag-of-Words (BOW) method, in which visual word occurrences alone are used to measure image similarity. Their approach employs a term-frequency inverse document-frequency (tf-idf) weighting scheme similar to that used in text retrieval. In current BOW methods, all descriptors from the initial image set are quantized and discarded while their geometric data are preserved for later matching. Quantization significantly reduces the storage requirements for features as invariant descriptors do not need to be retained, but can be summarized by a single cluster center for all features in a visual word. However, other information must still be retained for each feature, such as its source image ID, as well as location, scale, and orientation within that image for final geometric checking. In practice, the use of visual words provides, at most, a one order of magnitude reduction in memory usage regardless of the number of features within each visual word.

This section will discuss the properties of the ideal local feature, the different types of corners and region detectors, and different local invariant descriptors.

### 2.1.2 Properties of the Ideal Local Feature

Local features typically have a spatial extent, i.e., the local neighborhood of pixels mentioned above. In contrast to classical segmentation, this can be any subset of an image. The region boundaries do not have to correspond to changes in image appearance such as color or texture. Also, multiple regions may overlap, and uninteresting parts of the image such as homogeneous areas can remain uncovered.

Ideally, one would like such local features to correspond to semantically meaningful object parts. In practice, however, this is unfeasible, as this would

require high-level interpretation of the scene content, which is not available at this early stage. Instead, detectors select local features directly based on the underlying intensity patterns. Good features should have the following properties:

- **Repeatability:** Given two images of the same object under different viewing conditions, the features detected on the scene part visible in both images should be found in both images.
- **Distinctiveness/informativeness:** The intensity patterns underlying the detected features should show a lot of variation.
- **Locality:** The features should be local, so as to reduce the probability of occlusion and to allow simple model approximations of the geometric and photometric deformations.
- **Quantity:** The number of detected features should be a reasonable number. However, the optimal number of features depends on the application. Ideally, the number of detected features should be adaptable over a large range by a simple and intuitive threshold.
- **Accuracy:** The detected features should be accurately localized in image location, as with respect to scale and possibly shape.
- **Efficiency:** Preferably, the detection of features in a new image should allow for time-critical applications.
- **Invariance:** When large deformations are to be expected, the preferred approach is to model these mathematically if possible, and then develop methods for feature detection that are unaffected by these mathematical transformations.
- **Robustness:** In the case of relatively small deformations, it often suffices to make feature detection methods less sensitive to such deformations. Typical

deformations that are tackled using robustness are image noise, discretization effects, compression artifacts, blur, etc.

### 2.1.3 Local Feature Detectors

A large number of local feature detectors have been proposed in the literature. Example of local feature detectors are:

- Corner detectors:

Harris detector [45].

Smallest Univalue Segment Assimilating Nucleus (SUSAN) Detector [46].

Harris-Laplace: scale invariant corner detector [47].

Harris-Affine: affine-invariant corner detector [48].

- Blob detectors [49]:

Hessian detector.

Hessian-Laplace.

Hessian-Affine.

Difference-of-Gaussians (DoG) (used in Scale Invariant Feature Transform (SIFT) [50], (ASIFT) [51], (C-SIFT) [52], and (n-SFIT) [53]).

FAST: Features from Accelerated Segment Test.

- Region detectors:

Edge-based regions detector.

Salient region detector [54].

Intensity-based regions.

Maximally Stable Extremal Regions (MSER) [55].

## 2.1.4 Local Feature Descriptors

A large number of Local feature descriptors have been proposed in the literature. Example of local feature detectors are:

- SIFT [50].
- Shape context [56].
- SURF: Speeded Up Robust Features [57].
- DAISY descriptor [58].
- Binary descriptors:

LBP: Local Binary Patterns [59].

BRIEF: Binary Robust Independent Elementary Features [60].

ORB: Oriented FAST and Rotated BRIEF [61].

BRISK: Binary Robust Invariant Scalable Keypoints [62].

LIOP: Local Intensity Order Pattern for Feature Description [63].

Almost all of these descriptors have been extended from 2D to 3D. Details about these algorithms will not be discussed here because this is not the focus of the current research.

## 2.2 3D Detectors and Descriptors

The problem of 3D Shape Retrieval has become an active research area and attracted more researchers from several research communities including pattern recognition [9, 10], computer graphics [13], computer vision [64, 65], and applied mathematics [11].

When try to adapt feature-based approaches applied to 2D images to 3D shapes, have to consider that the invariance type in non-rigid shapes is different

than these in images, where variation of transformations is much larger, including changes in pose, bending, and connectivity changes. Since many natural shape deformations such as bending can be approximated by isometries, so the shape descriptor should depend on intrinsic properties of the shape, this dependency makes it invariant to such deformations.

Also shapes are not rich in features like images, making it harder to detect a large number of repeatable and stable keypoints. This motivates a lot of researchers to avoid feature detection at all and use dense descriptors instead. As such, the area of 3D shape descriptors for shape matching and retrieval still needs more work.

Although literature is rich with several shape descriptors definitions, they can be classified under two main categories: model-based or view-based. Model based methods directly make use of the 3D data, while view based methods store a number of 2D projections of the 3D object. Previous investigations, such as [15], [16] demonstrate that view-based methods with pose normalization pre-processing can achieve better performance in retrieving rigid shapes, but at higher storage and computational cost than many other 3D descriptors.

Model based techniques can be categorized into global feature-based techniques and local feature techniques. Local matching [66] is performed between sparse points or regions of the shape, while in the global shape approaches similarity is measured among entire models [67]. For global or whole-to-whole shape matching, various descriptors have been proposed for representing the distribution of a shape's surface area. The Spherical Harmonic descriptor [68] represents the distribution of surface area in each shell as a series of spherical harmonic coefficients. Other descriptors include geometric moments [69], distribution of pair-wise Euclidean distances [70], volume and area descriptors [71], Self-similarity (symmetry) Kazhdan et al. [68], and the work by Hassouna et al. [72] which represents a 3D shape by a set of one-dimensional curves that are locally symmetric with



respect to its boundary. Evaluation of 3D shape retrieval methods with respect to other several requirements can be found in [12].

Although global shape descriptors have shown good performance on many data sets, they have an underlying assumption that shapes are rigidly transformed. Different approaches used local feature detection and a local descriptor to describe 3D shapes, such as spin images [19], surface point signature [20], contour and edge detection features [21], local patches [22], and conformal factor [23]. However these methods cannot deal with the non-rigid shape deformation, and can not cover the properties of the desired shape descriptor.

The problem of non-rigid shape deformation needs more work to compensate for the degrees of freedom resulting from local deformations. Early work by Elad and Kimmel [4, 24] proposed modeling shapes as metric spaces with the geodesic distances as an intrinsic metric, these spaces are invariant to inelastic deformations. Ling and Jacobs [25] and Bronstein et al. [26] used this framework with a metric defined by internal distances in 2D shapes. Reuter et al. [2, 27] used the Laplacian spectra as intrinsic shape descriptors, and they employed the Laplace-Beltrami spectra as 'shape-DNA' or a numerical fingerprint of any 2D or 3D manifold (surface or solid). They proved that 'shape-DNA' is an isometry-invariant shape descriptor. Rustamov [28] used an isometry-invariant shape representation in the Euclidean space, and then histograms of Euclidean distances to compare between shapes.

Another type of intrinsic geometry is generated by heat diffusion processes on the shape. Coifman and Lafon [73] popularized the notation of diffusion geometry, which is closely related to scale-space methods in image processing. Recently Sun et al. [74] proposed heat kernel signatures (HKS) as a deformation-invariant descriptor based on diffusion of multi-scale heat kernels. HKS is a point based signature satisfying all of the good descriptor properties except for scale invariance. It characterizes each vertex on the meshed surface using a vector. However,

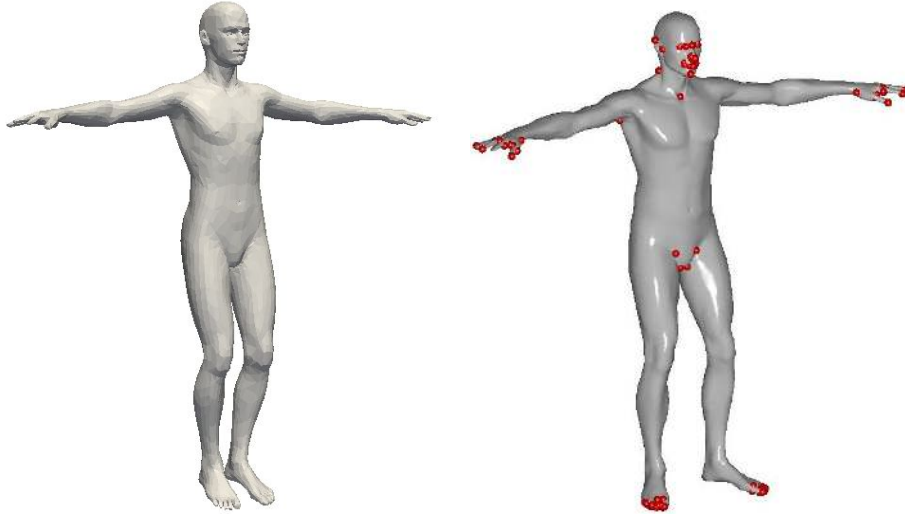


Figure 2.2: Example of a 3D shape to the left and the detected points to the right using Harris 3D detector.

the authors did not demonstrate how to retrieve shapes using HKS, although they pointed out its potential in shape retrieval applications. Bronstein et al. [75] tried to tackle the scale problem in HKS. Fang et al. [1, 76] defined the temperature distribution (TD) of the heat mean signature (HMS) as a shape descriptor for shape matching, and they used the same heat kernel for mesh segmentation. Their TD is a global shape descriptor and they used the  $L2$  norm which is a very basic matching method to compute the distance between two TD descriptors. Shape Google work by Bronstein et al. [3] introduces a shape retrieval method based on the Heat Kernel Signature (HKS) and its scale-invariant version. The idea is to evaluate HKS at all points of a shape or alternatively at some feature points of a shape and to represent the shape by a Bag of Features (BoF) vector. Sparsity in the time domain is enforced by preselecting some values of the time  $t$ .

There are several local detection and description algorithms for 3D shapes. An overview some of them follows.

### 2.2.1 3D Feature Detectors

An interest point detector in 3D mesh must have some of these properties:

- It must be invariant to affine transformations.
- It must be robust to noise, which can be introduced during the capture process.
- It must be robust to different tessellations.

This section will examine some examples of the feature detectors.

### 2.2.1.1 Harris 3D

The Harris operator is an effective feature detection method, first proposed in images [45], and was extended to 3D shapes by Sipiran and Bustos [77]. This method detects interest points for 3D objects based on the Harris operator. They propose an adaptive technique to determine the neighborhood of a vertex, over which the Harris response on that vertex is calculated. Figure 2.2 shows an example of detected points using this approach.

### 2.2.1.2 Mesh DOG

Difference of Gaussians (DOG) is a classical feature detection approach used in computer vision. Among these methods that use the DOG for feature detection in 3D is Zaharescu et al. [78] which introduces the mesh DOG approach by first applying Gaussian filtering to functions defined on the shape. The extrema of the functions Laplacian (DOG) are found across scales using a one-ring neighbourhood, then detected extrema are thresholded. After that the unstable extrema are eliminated. Examples of detected features using Mesh DOG are shown in Figure 2.3.

### 2.2.1.3 Heat Kernel Feature Detectors

Sun et al. [74] proposed a feature detection methods based on the heat kernel. These methods define a function on the shape, and measure the amount

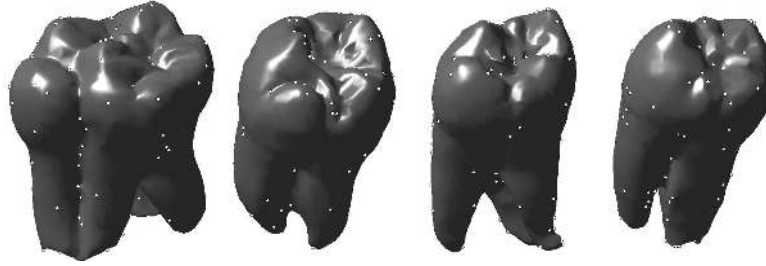


Figure 2.3: Examples of detected features using Mesh DOG on individual 3D human teeth.

of heat remaining at a point  $x$  after large time  $t$  given a point source at  $x$  at time 0. Then use local maxima of this function as feature points.

## 2.2.2 3D Feature Descriptors

A local descriptor for a 3D shape must have some of these properties defined in Table 1.1. Some examples of local description algorithms will be examined below.

### 2.2.2.1 Shape Context

Though originally proposed for 2D shapes and images [56], shape context has also been generalized to 3D shapes. Such a descriptor is translation-invariant and can be made rotation-invariant. It is computable on any kind of shape representation, including point clouds, voxels, and triangular meshes, but it is not deformation invariant, therefore it is not suitable for non-rigid shape analysis applications.

### 2.2.2.2 Spin Images

Perhaps one of the best known classes of feature descriptors are spin images, which is introduced by Johnson and Hebert [19]. This method originally developed for rigid shape comparison, and is thus very sensitive to non-rigid shape deformations.

### 2.2.2.3 Mesh HOG

Zaharescu et al. [78] use the histogram of gradients of a function defined in a local neighborhood of a point, as a point descriptor (similar to the Histogram of Gradients (HOG) technique used in computer vision). Zaharescu et al. [78] show insensitivity of their descriptor to non-rigid deformations.

### 2.2.2.4 Geodesic Distance

A Riemannian manifold is a manifold  $\mathcal{M}$  with a Riemannian metric  $U(u, v)$ , also called the first fundamental form. The Riemannian metric is an inner product in the tangent space and can be defined for two tangent vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  with the help of the first fundamental matrix  $G$ . After calculating the lengths, the shortest distances can be computed between two points  $\mathbf{x}, \mathbf{y} \in \mathcal{M}$  along the surface. This shortest distance called the geodesic distance  $D_{\mathbf{y}}(\mathbf{x})$ , between any point  $\mathbf{x}$  and  $\mathbf{y}$  can be obtained by solving the Eikonal equation

$$|\nabla_{T_{\mathbf{x}_i}} D(\mathbf{x})| = 1, \quad (2.1)$$

with boundary condition  $D_{\mathbf{y}}(\mathbf{y}) = 0$ .

### 2.2.2.5 ShapeDNA

The ShapeDNA has been introduced in [2] as the first spectral method used for non-rigid shape analysis. Spectral methods have later been employed by the authors for local shape analysis of structures in the human brain to analyse disease effects [27] and for automatic shape segmentation [6]. ShapeDNA is the first  $n$  eigenvalues of the Laplace-Beltrami Operator (LBO) for 2D surfaces or 3D solids. The proposed approach will be compared with ShapeDNA in Chapter 3.

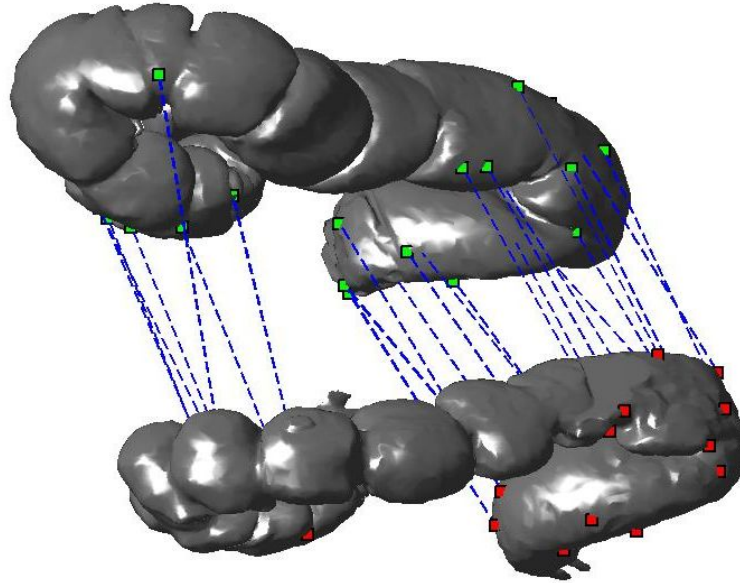


Figure 2.4: Example of feature matching using meshSIFT on 3D colon parts.

#### 2.2.2.6 Heat Kernel Signatures (HKS)

Recently, there has been increased interest in the use of diffusion geometry for shape recognition by Rustamov [28], Sun et al. [74] Bronstein et al. [75], and Raviv et al. [79]. This type of geometry arises from the heat equation.

#### 2.2.2.7 Scale Invariant Heat Kernel Signatures (SI-HKS)

Bronstein et al. [75] solved the HKS scale problem through a series of transformations. The same research group has recently introduced the Shape Google approach [3] based on the scaled-invariant HKS. The idea is to use HKS at all points of a shape, or alternatively at some shape feature points, to represent the shape by a Bag of Features (BoF) vector. Chapter 3 will compare the proposed approach with the technique used in Shape Google [3].

### **2.2.2.8 Scale Invariant Feature Transform for Meshes (meshSIFT)**

The meshSIFT [80] algorithm is an extension of the 2D SIFT [36] to 3D meshes. It consists of four major components: keypoint detection, orientation assignment, the local feature description and feature matching. An example of feature matching using meshSIFT is shown in Figure 2.4.

## **2.3 Summary**

This chapter reviewed the existing 2D and 3D descriptors, and gave a quick literature review about the 2D descriptors. Then it introduced the global and local 3D shape descriptors, also this chapter gave some examples of the famous 3D feature detectors and descriptors.

## CHAPTER 3

# FUNDAMENTALS OF HEAT EQUATION AND THE LAPLACE-BELTRAMI

It is well-known that heat diffusion on the surface of a shape is fully described by the heat kernel, associated with the Laplace-Beltrami operator. This chapter, will study the solution of the heat equation and the reconstruction of the heat kernels (HK) as a promising direction for shape representation. It starts with an introduction to fundamentals of geometry, topology, the length spaces, isometries, manifolds, embedded surfaces, first fundamental form, curvature and the second fundamental form that will allow for the formulation of properties of non-rigid objects.

The chapter, will detail the derivation of the heat equation, describe the Initial Boundary Value Problem (IBV), give the details of the the Laplace-Beltrami operator, its properties, and the discretization of the Laplace-Beltrami operator and how to use it to solve the heat equation IBV problem on a manifold. Then it will talk about the solution of the heat equation and the basics of diffusion on Riemannian manifolds that are necessary to define the heat kernel, and how to solve for the heat kernels using Finite Element Method (FEM).

### 3.1 Fundamentals of Geometry

Geometry means earth measure, this ancient term used by greek Eratosthenes "who first measured the earth". He simultaneously (by hourglasses) mea-



sured shadows at two locations in Egypt to calculate the distance between them, and so, the earth's circumference was roughly extrapolated. Also ancient Egyptian geometers knew that the value of pi was represented by the fraction  $22/7$ , so 880 royal cubits was the length of two base sides of the Great Pyramid, and 280 royal cubits for its height, because  $880/280$  equals pi.

Euclid and Archimedes then significantly advance the field of the modern geometry. Today, geometry is very important for engineers, they use it for modeling of physical objects. Geometric tools are necessary for the description of the non-rigid shapes. This section, will give an introduction for geometry fundamentals that will allow to formulate properties of non-rigid objects.

### 3.1.1 Basic Terms in Geometry and Topology

Distance is one of the most fundamental concepts in geometry. The distance between two points can be measured as the length of the straight line connecting them. However, the three-dimensional space with the Euclidean distance is simply a particular instance of a more general notion called the metric space. Before defining the distance the function and what is the meaning of continuity need to be defined.

**Definition 5 *Function:*** A **function**  $f$  is a well-defined rule assigning to each element of a set  $A$  a unique element in the set  $B$  it is denoted by  $f : A \rightarrow B$ . The set  $A$  is the **domain** of the function  $f$  and the receiving set  $B$  is its **codomain**. The **image** or **range** of the function can be defined as  $f(A) = \{f(a) \in B \mid a \in A\}$  which is a subset of  $B$ .

**Definition 6 *Continuous Function:***  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a single-variable function, where  $\mathbb{R}$  is the set of real numbers, this function is continuous at a point  $x_o \in \mathbb{R}$  if for every  $\varepsilon > 0$ , there is  $\delta > 0$ , such that whenever  $|x - x_o| < \delta$ , then  $|f(x) - f(x_o)| < \varepsilon$ .

Then the metric space need to defined.

**Definition 7 Metric Space:** A *metric space* is a set  $\mathbf{X}$  together with a distance function  $d : \mathbf{X} \times \mathbf{X} \rightarrow \mathbb{R}$  satisfying the following:

1.  $d(x, y) \geq 0$  for all  $x, y \in \mathbf{X}$ .
2.  $d(x, y) = 0$  if and only if  $x = y$
3.  $d(x, y) = d(y, x)$  for all  $x, y \in \mathbf{X}$ .
4. The triangle inequality:  $d(x, y) + d(y, z) \geq d(x, z)$  for all  $x, y, z \in \mathbf{X}$ .

In topology, the most general object to deal with is a set of points on which functions can be defined. This sub-section, will investigate sets of points at a microscopic level, beginning with definitions and notations.

Consider point sets which are subsets of the real Euclidean  $n$ -space which is defined as follows: Real Euclidean  $n$ -space is given by  $\mathbb{R}^n = \{\mathbf{x} = (x_1, x_2, \dots, x_n) | x_i \in \mathbb{R}\}$  where  $\mathbf{x}$  denotes a point with  $n$ -coordinates. A distance measure for Euclidean space can be defined using the Euclidean metric which is given by the following definition. Given two points  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ , the Euclidean distance (metric) between  $\mathbf{x}$  and  $\mathbf{y}$  is given by:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

Hence, a disc or ball centered at  $x$  with radius  $r$  can be defined as  $D^n(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < r\}$ .  $D^n(\mathbf{x}, r)$  is called (disc) **neighborhood** of  $\mathbf{x} \in \mathbb{R}^n$  which is an open disc containing  $x$ . Consider a set of points in the Euclidean  $n$ -space i.e.  $\mathbb{R}^n$ , a point  $\mathbf{x}$  can be related to this set in terms of its neighborhood in one of the following ways.

Let  $A$  be a set of points with  $A \subseteq \mathbb{R}^n$  and  $\mathbb{R}^n - A$  be all points not in  $A$ , i.e. the complement of  $A$ :

1. A point  $\mathbf{x} \in \mathbb{R}^n$  is an **interior** point of  $A$  if there is a neighborhood  $\mathcal{N}$  of  $\mathbf{x}$  such that  $\mathbf{x} \in \mathcal{N} \subseteq A$ , i.e. the disc is totally enclosed in  $A$ .

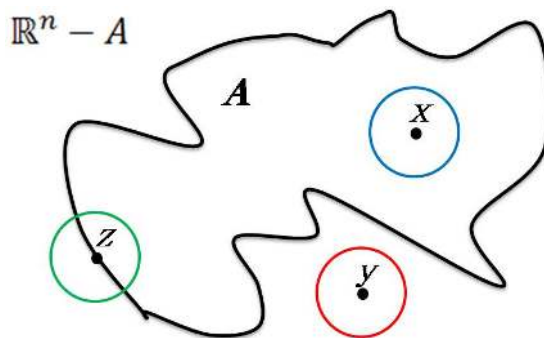


Figure 3.1: Point  $\mathbf{x}$  is an interior point of  $A$ , point  $\mathbf{y}$  is an exterior point of  $A$ , while  $\mathbf{z}$  is a boundary point of  $A$ .

2. A point  $\mathbf{x} \in \mathbb{R}^n$  is an **exterior** point of  $A$  if there is a neighborhood  $\mathcal{N}$  of  $\mathbf{x}$  such that  $\mathbf{x} \in N \subseteq (\mathbb{R}^n - A)$ , i.e. the disc is totally outside  $A$ . Another way of saying this is  $N \cap A = \phi$ .
3. A point  $\mathbf{x} \in \mathbb{R}^n$  is a **limit/accumulation** point of  $A$  if every neighborhood  $\mathcal{N}$  of  $\mathbf{x}$  contains at least one point of  $A$ , thus  $N \cap A \neq \phi$ . Hence every point of  $A$  is a limit point of  $A$ .
4. A point  $\mathbf{x} \in \mathbb{R}^n$  is an **isolated** point of  $A$  if  $\mathbf{x} \in A$  and it has a neighborhood  $N$  satisfying  $N \cap A = \{\mathbf{x}\}$ . Hence a point  $\mathbf{x}$  is not isolated if every neighborhood of  $x$  contains at least one point in  $A$  other than itself.
5. A point  $\mathbf{x} \in \mathbb{R}^n$  is a **boundary/frontier** point of  $A$  if every neighborhood  $\mathcal{N}$  of  $\mathbf{x}$  intersects both  $A$  and  $\mathbb{R}^n - A$ , i.e. contains points in and outside  $A$ . Thus boundary points are also limit points.

see Figure 3.1 for the definition of interior, exterior, and boundary points.

It is important to note that any point in  $\mathbf{x} \in \mathbb{R}^n$  must be either interior, exterior or a boundary point of  $A$ , a point cannot be both interior and exterior to  $A$ . Also note that interior points are elements of  $A$  while exterior points are elements of its complement, on the other hand boundary and limit points might lie in  $A$  or its complement.

The definition of open and closed sets will be as following:

**Definition 8 *Open set:*** A set  $A$  is said to be **open** if every point  $\mathbf{x} \in A$  is an interior point.

**Definition 9 *Closed set:*** A set  $A$  is said to be **closed** if every point  $\mathbf{x} \notin A$  is an exterior point.

Another way to view an open set is that it does not contain points on its boundary hence all points which lie in an open set are interior points to this set. However, if a set is not open this does not mean that it is closed, a set might neither be open nor closed. On the other hand a closed set means that any point lying outside this set is totally outside it, i.e. points that are not in a closed set are neither interior nor boundary points to this set.

**Definition 10 *Bounded set:*** A set  $A \subseteq \mathbb{R}^n$  is said to be **bounded** if  $A \subseteq D^n(0, r)$  for some  $r$ . Thus the set  $A$  can be enclosed in some sufficiently large disc centered at its origin, i.e.  $A$  does not go on forever.

**Definition 11 *Connected set:*** A set  $A$  is (sequentially) **compact** if every infinite sequence of points in  $A$  has a limit point in  $A$ , that is if  $\{\mathbf{x}_i\}_{i=1}^{\infty}$  is a sequence and  $\mathbf{x}_i \in A$  for each  $i$ , then there is a point  $\mathbf{x} \in A$  such that  $\mathbf{x}$  is a limit point of  $\{\mathbf{x}_i\}_{i=1}^{\infty}$ .

Note that this definition has two claims, the first one is that every sequence has at least one limit points and the second is that this limit point will actually lie inside the set. Consider the following examples.

**Examples:**

1. The real line  $\mathbb{R}$  is not compact, since the sequence  $\{1, 2, 3, \dots\} = \{n\}_{n=1}^{\infty}$  consists of points in  $\mathbb{R}$  but has no limit points in  $\mathbb{R}$ .

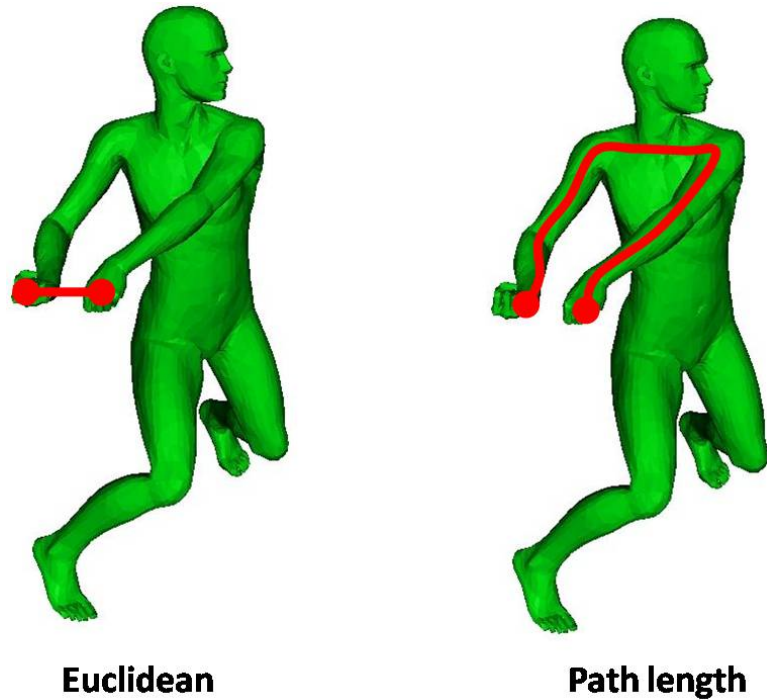


Figure 3.2: Two types of metrics.

2. The interval  $(0, 1)$ , the number 0 is a limit point for the sequence  $\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\} = \{\frac{1}{n}\}_{n=2}^{\infty} \subseteq (0, 1)$ , but  $0 \notin (0, 1)$ , hence the open interval  $(0, 1)$  is not compact.

Another fundamental notion in topology is the number of pieces or components an object has, if an object contains only one piece, it is considered connected; this is true if all its parts are stuck to each other.

**Definition 12 Connected set:** A set  $S$  is **connected** if whenever  $S$  is divided into two non-empty sets such that  $S = A \cup B$ ,  $A \neq \emptyset$ ,  $B \neq \emptyset$  and  $A \cap B = \emptyset$ , then either  $A$  or  $B$  contains a limit point of the other.

### 3.1.2 Length Spaces

The Metric spaces is defined in Section 3.1.1, metric spaces allow the definition of an abstract distance function between points. However, in many cases the notion of distance is somehow unclear. The example given by Burago et

al. [81] illustrate the difference between the length of the path and the Euclidean distance. Using the Euclidean distance, one can claim that the distance between New York and Sydney is about 8,000 miles; however, this number is of little use to an aircraft pilot, as it assumes travelling along a straight tunnel dug through the Earth between the two cities. The meaning of distance is born from the lengths of paths. Now Consider the path as a continuous map  $\gamma : [a, b] \rightarrow \mathbf{X}$  of an interval. And let length of the path to be the function  $L(\gamma)$  that assigns a nonnegative number to every path. Using these two concepts, A metric induced by the length can be defined as

$$d_L(x, y) = \inf_{\gamma} \{L(\gamma) \quad s.t. \quad \gamma : [a, b] \rightarrow X, \gamma(a) = x1, \gamma(b) = x2\} \quad (3.1)$$

The distance between two points is thereby the infimum of lengths of paths connecting between them. Such a metric is called a length metric and a metric space  $(\mathbf{X}, d_L)$  is called a length space. See Figure 3.2 for the difference between the Euclidean distance and the path length.

### 3.1.3 Manifolds

**Definition 13 *Manifold:*** An  $n$ -dimensional **manifold** is a topological space such that every point has a neighborhood topologically equivalent to an  $n$ -dimensional open disc with center  $\mathbf{x}$  and radius  $r$  i.e.  $D^n(\mathbf{x}, r) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{x}| < r\}$ . It is required that any two distinct points have disjoint neighborhoods. A 2-manifold is often called a **surface**.

There may exist metric and length spaces that are not manifolds and manifolds without a metric. This discussion will focus on manifolds that are also metric or length spaces. Also to distinguish between a manifold and a manifold with a boundary (see Figure 3.3). Consider as an example a closed subset  $A \subset \mathbf{R}^n$  of the

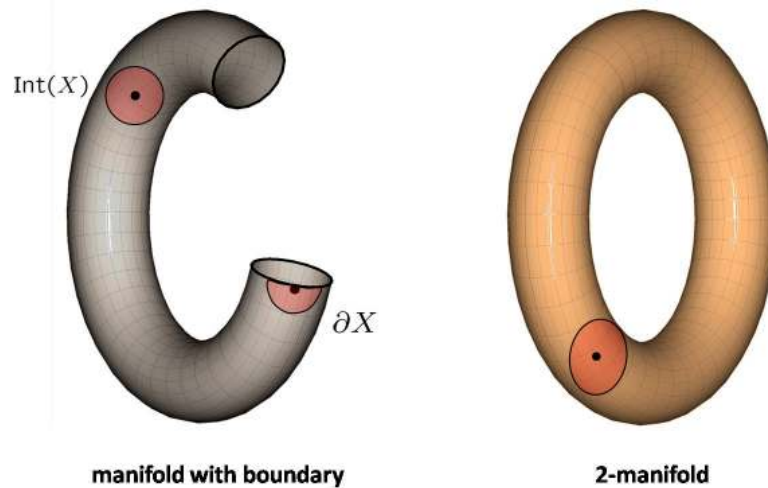


Figure 3.3: The two types of manifolds.

Euclidean space. Points inside  $A$  have open neighborhoods homeomorphic to  $\mathbf{R}^n$  and thus the interior of  $A$  is, indeed, a manifold. However, points belonging to the "edge" of  $A$  are clearly not homeomorphic to  $\mathbf{R}^n$  but to the closed Euclidean half-space  $[0, \infty) \times \mathbf{R}^{n-1}$ . So a manifold with a boundary is a non-empty space  $X$  in which every point has a neighborhood that can be charted either in  $\mathbf{R}^n$  or in  $[0, \infty) \times \mathbf{R}^{n-1}$ . The set of points in  $X$  that can be charted only in  $[0, \infty) \times \mathbf{R}^{n-1}$  is called the boundary and denoted by  $\partial X$ , whereas their complement is referred to as the interior and is denoted by  $int(X)$ .

**Example:** The sphere, denoted by  $\mathbb{S}^2$ , is a surface, even though it exists in 3-dimensional space (this emphasizes the difference between intrinsic and extrinsic properties, where intrinsic properties have to do with the object itself, in contrast to extrinsic properties which describe how the object is embedded in the surrounding space). If one considers a point  $\mathbf{x} \in \mathbb{S}^2$  as a point in  $\mathbb{R}^3$ , it will have a neighborhood that looks like a ball. However as a point on the sphere, with the relative topology,  $\mathbf{x}$  has neighborhoods of the form:  $\mathcal{N} = \mathcal{D}^n(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^3 : \|\mathbf{x}-\mathbf{y}\| < r \text{ and } \mathbf{y} \in \mathbb{S}^2\}$ , these neighborhoods look like 2-dimensional discs which have been warped a bit.

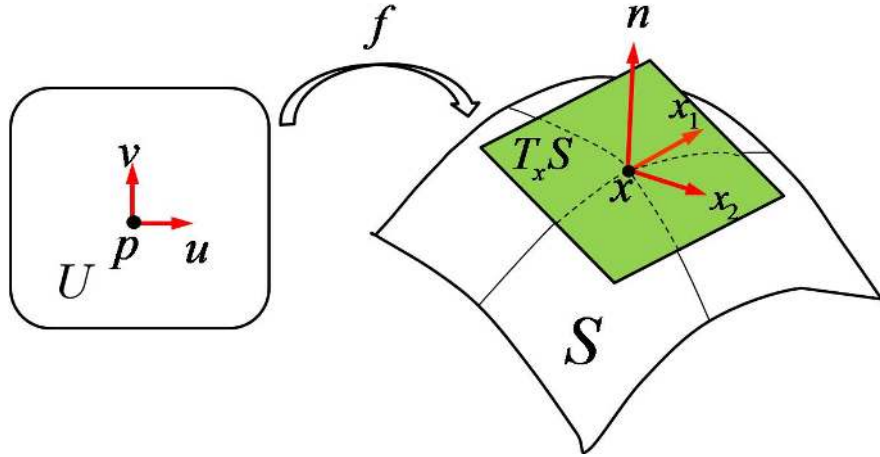


Figure 3.4: 3D surface parametrization, tangent plane, and normal.

### 3.1.4 Embedded Surfaces

Surfaces of all physical objects can be represented as smooth two-dimensional manifolds in three-dimensional Euclidean space. Such manifolds are often referred to as embedded surfaces. Such surfaces can often be described by a smooth map  $f : U \rightarrow \mathbb{R}^3$  from a subset  $U$  of  $\mathbb{R}^2$  to  $\mathbb{R}^3$ . The set  $U$  is called a parametrization domain. The parametrization  $f$  maps all points  $\mathbf{p} = (u, v) \in U$  to a point  $\mathbf{x} = f(u, v)$  on the 3D surface. And the surface  $\mathbf{S}$  which has a global parametrization (i.e. a local parametrization  $(U, f)$  for which  $f(U) = \mathbf{S}$ ) is called a simple surface.

If  $\mathbf{S}$  is a surface and  $(U, f)$  is a local parametrization of  $\mathbf{S}$ , then, if  $\mathbf{x}(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v))$ , then the equations  $x_1 = x_1(u, v)$ ,  $x_2 = x_2(u, v)$ ,  $x_3 = x_3(u, v)$ , where  $(u, v) \in U$ , are called the parametric equations of the surface.

When the derivatives  $\mathbf{x}_1 = \partial_u \mathbf{x}$  and  $\mathbf{x}_2 = \partial_v \mathbf{x}$  of  $\mathbf{x}$  with respect to the coordinates are linearly independent for every  $(u, v) \in U$ , then  $x$  is regular. In such a case, the vectors  $\mathbf{x}_1, \mathbf{x}_2$  span a two-dimensional space at  $\mathbf{x} = f(u, v)$ , referred to as the tangent plane or the tangent space and denoted by  $T_{\mathbf{x}}\mathbf{S}$ . the vector perpendicular to the tangent plane is called the normal to the surface and is denoted by  $\mathbf{n}$ . See Figure 3.4.



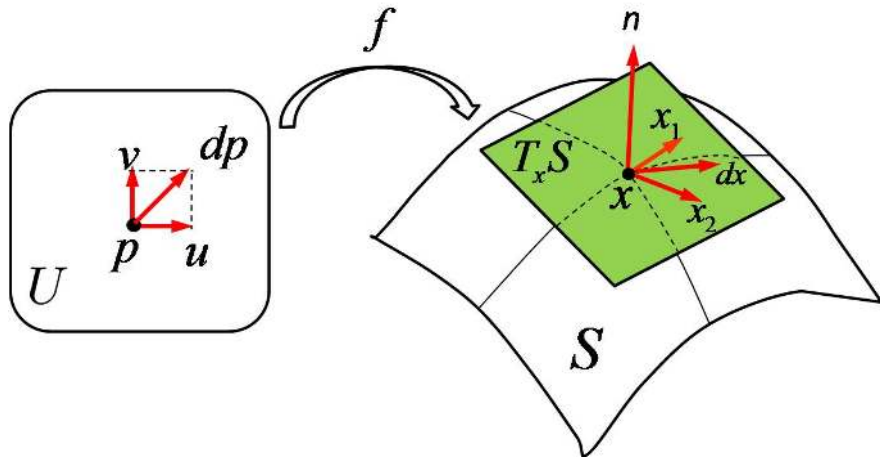


Figure 3.5: Illustration of first fundamental form.

### 3.1.5 First Fundamental Form

Let  $\mathbf{p}$  be a point in the parametrization domain  $U$  and the corresponding point  $\mathbf{x} = f(\mathbf{p})$  on the surface  $\mathbf{S}$ , and consider a displacement by  $d\mathbf{p} = (du, dv)$  around  $\mathbf{p}$  will displace the point on the surface to  $f(\mathbf{p} + d\mathbf{p}) = \mathbf{x} + \mathbf{x}_1 du + \mathbf{x}_2 dv = \mathbf{x} + Jd\mathbf{p}$ . Here,  $J$  is a  $3 \times 2$  matrix having  $\mathbf{x}_1$  and  $\mathbf{x}_2$  as the columns; this matrix is called the Jacobian of the parametrization  $f : U \rightarrow \mathbf{S}$  at a point  $\mathbf{p}$ . To quantify the length of the displacement  $d\mathbf{x} = Jd\mathbf{p}$ , it can be written as:

$$dl^2 = \|d\mathbf{x}\|^2 = \|Jd\mathbf{p}\|^2 = d\mathbf{p}^T J^T J d\mathbf{p} = d\mathbf{p}^T G d\mathbf{p} \quad (3.2)$$

where  $G = J^T J$  is a symmetric  $2 \times 2$  matrix, whose elements are the inner products  $g_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ . And  $G$  is positive definite since  $g_{11}g_{22} - g_{21}^2 > 0$ . The quadratic form 3.2 is called the **first fundamental form** of the surface or the **Riemannian metric**. See Figure 3.5. The first fundamental form of a surface completely describes its intrinsic geometry or the length metric of the surface.

### 3.1.6 Curvature and the Second Fundamental Form

The curvature can be expressed in terms of how fast the normal vector rotates with the movement on a surface. The directional derivative of the normal vector can be used to measure the rate of change of the normal. Given a point  $\mathbf{x} \in T_{\mathbf{x}}\mathbf{S}$  and a direction  $\mathbf{v} \in T_{\mathbf{x}}\mathbf{S}$ , the directional derivative of  $N$  is defined as

$$D_{\mathbf{v}}N = \lim_{t \rightarrow 0} \frac{1}{t} (N(\gamma(t)) - N(\mathbf{x})) = \left. \frac{d}{dt} N(\gamma(t)) \right|_{t=0} \quad (3.3)$$

where  $\gamma$  is a path or smooth curve passes through  $\mathbf{x}$  and  $\mathbf{v}$ .  $D_{\mathbf{v}}N \in T_{\mathbf{x}}\mathbf{S}$  is a vector in  $\mathbb{R}^3$  measuring the change in  $N$  as making a differential step in the direction  $\mathbf{v}$ , and it is perpendicular to the normal. The second fundamental form can be defined as the matrix  $B$  with elements  $b_{ij} = -\langle \partial_{u^i} N, \mathbf{x}_j \rangle$ , where  $u^i \in \{u, v\}$ . the second fundamental form is responsible for the extrinsic geometry of the surface, that is, the way the surface resides in the Euclidean space.

Now for any point  $\mathbf{x} \in \mathbf{S}$  let  $E = \langle \mathbf{x}_1, \mathbf{x}_1 \rangle$ ,  $F = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$ , and  $G = \langle \mathbf{x}_2, \mathbf{x}_2 \rangle$

The unit surface normal  $\mathbf{n}$  can be calculated as:

$$\mathbf{n} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{\|\mathbf{x}_1 \times \mathbf{x}_2\|} \quad (3.4)$$

where,  $\times$  denotes the cross product.

The second fundamental forms  $L$ ,  $M$  and  $N$  of  $S$  are expressed as follows.

$$L = \langle \mathbf{x}_{11}, \mathbf{n} \rangle, \quad M = \langle \mathbf{x}_{12}, \mathbf{n} \rangle, \quad N = \langle \mathbf{x}_{22}, \mathbf{n} \rangle,$$

where the second derivatives  $\mathbf{x}_{11} = \partial_u \partial_u \mathbf{x}$ ,  $\mathbf{x}_{12} = \partial_u \partial_v \mathbf{x}$  and  $\mathbf{x}_{22} = \partial_v \partial_v \mathbf{x}$  of  $\mathbf{x}$  with respect to the coordinates are linearly independent for every  $(u, v) \in U$ . Then the maximum principal curvature  $\kappa_1$  and the minimum principal curvature

$\kappa_2$  are defined as the two roots  $\lambda_1$  and  $\lambda_2$ , which satisfy the following identity:

$$\begin{vmatrix} E\lambda - L & F\lambda - M \\ F\lambda - M & G\lambda - N \end{vmatrix} = 0 \quad (3.5)$$

Simplifying Equation 3.5, then:

$$(EG - F^2)\lambda^2 + (2FM - EN - GL)\lambda + (LN - M^2) = 0 \quad (3.6)$$

As a result, Gaussian curvature  $K$  and mean curvature  $H$  can be defined as:

$$K = \kappa_1\kappa_2 = \lambda_1\lambda_2 = \frac{LN - M^2}{EG - F^2} \quad (3.7)$$

$$H = \frac{1}{2}(\kappa_1 + \kappa_2) = \frac{1}{2}(\lambda_1 + \lambda_2) = \frac{EN - 2FM + GL}{2(EG - F^2)} \quad (3.8)$$

Then the two principal curvatures can be computed as

$$\kappa_1 = H + \sqrt{H^2 - K} \quad (3.9)$$

$$\kappa_2 = H - \sqrt{H^2 - K} \quad (3.10)$$

The values of  $K$  and  $H$  define the local behavior of the surface, that is, how the surface is curved at the point  $\mathbf{x}$ . For example, a plane has  $K = H = 0$  at every point. A sphere has  $K > 0$ . A hyperbolic surface has  $K < 0$  because one of the principal curvatures is positive and one is negative. See Figure 3.6.

### 3.1.7 Isometries

Isometries copy metric geometries which make isometric spaces equivalent from the point of view of metric geometry. Two metric spaces  $(\mathbf{X}, \alpha)$  and  $(\mathbf{Y}, \beta)$  are equivalent if there exists a distance-preserving map (isometry)  $\varphi : (\mathbf{X}, \alpha) \rightarrow (\mathbf{Y}, \beta)$  satisfying  $\beta \circ (\varphi \times \varphi) \approx \alpha$  such that  $(\mathbf{X}, \alpha)$  and  $(\mathbf{Y}, \beta)$  are called isometric,



Figure 3.6: illustration of maximum principal curvature  $\kappa_1$  and the minimum principal curvature  $\kappa_2$  in the upper row, and Gaussian curvature  $K$  and mean curvature  $H$  in the lower row.

and denoted by  $(\mathbf{X}, \alpha)(\mathbf{Y}, \beta)$ . Examples of isometries are shown in Figure 3.7. The three shapes are isometric with respect to  $\alpha$  the original shape metric is  $(\mathbf{X}, \alpha)$ , after some transformation (rotation, translation, or reflection) it gives the Euclidean isometric metric  $(\tau\mathbf{X}, \alpha)$ , then after deformation it gives the second isometric metric  $(\mathbf{X}, \beta)$ .

The previous section gave an introduction to the geometry fundamentals. It discussed the definition of basic terms in geometry and topology, the length spaces, isometries, manifolds, embedded surfaces, first fundamental form, curvature and the second fundamental form. These terminologies will be used in the following section, and it will allow to formulate the heat equation.

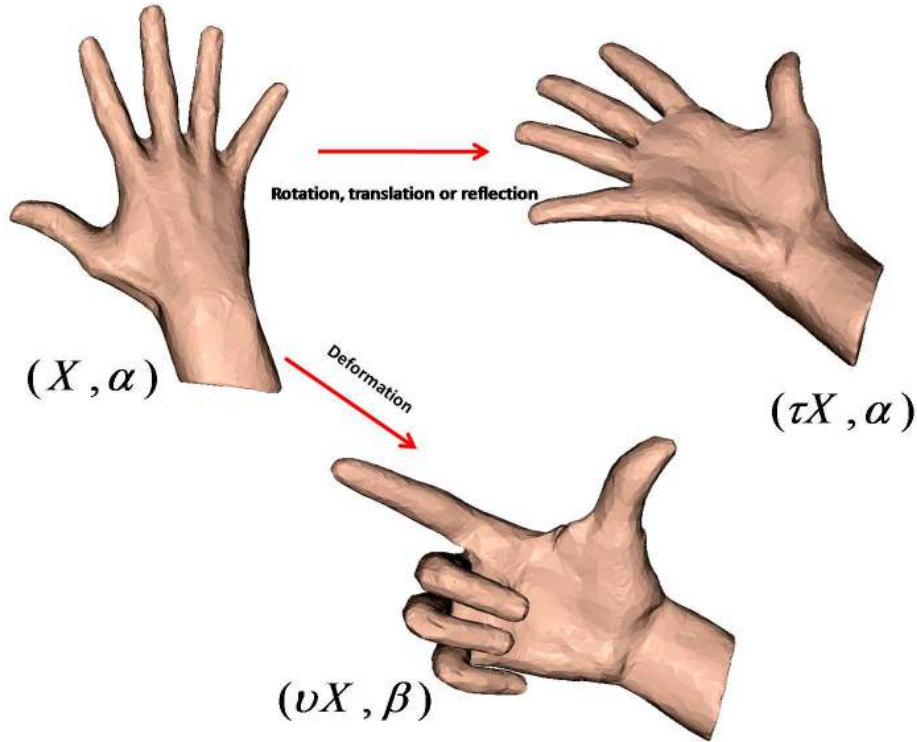


Figure 3.7: Two types of isometries, the original shape  $(\mathbf{X}, \alpha)$ , the Euclidean isometric copy  $(\tau\mathbf{X}, \alpha)$ , and the deformed isometric copy  $(\mathbf{X}, \beta)$ .

### 3.2 The Heat Equation

Heat<sup>1</sup> is energy transferred from one system to another by thermal interaction, it is always accompanied by a transfer of entropy. The heat equation is an important partial differential equation which describes the distribution of heat (or variation in temperature) in a given region over time. This section, will discuss the derivation of the heat equation from the physical properties of the object, and from the rate of heat flow through the boundaries of the object. Since the Laplace-Beltrami and the heat diffusion operator, together with the corresponding Laplacian eigenmaps have been used in several contexts for shape representation, the Initial Boundary Value Problems (IBV) will be described. Also the details of the the Laplace-Beltrami operator, its properties, the discretization of the Laplace-Beltrami operator will be given, and how Laplace-Beltrami operator

<sup>1</sup>In this discussion when mention the term heat it refer to the quantity of heat in an object.

will be used to solve the heat equation IBV problem on a manifold. After that this section will talk about the solution of the heat equation and the basics of diffusion on Riemannian manifolds that are necessary to define the heat kernel. Finally it will mention the heat kernel signature and its properties.

For the derivation of the heat equation in 1D see Appendix A. The heat equation, also known as the diffusion equation models the flow of heat at time  $t$ . The heat equation is a second order parabolic partial differential equation, and usually written as

$$U_t = kU_{xx} \tag{3.11}$$

### 3.2.1 Initial Boundary Value Problems (IBV)

In order to solve the heat equation the problem initial conditions should be assumed. The temperature of every point along the rod at time  $t = 0$  need to be defined, so the problem is to find  $U(x, t)$  such that

$$\begin{aligned} U_t(x, t) &= kU_{xx}(x, t) && \text{for } 0 \leq x \leq L \text{ and } t > 0, \\ U(x, 0) &= u(x) && \text{for } 0 \leq x \leq L, \\ U(0, t) &= T_0 && \text{and } U(L, t) = T_L \text{ for } t > 0 \end{aligned} \tag{3.12}$$

This is the general form for the heat equation as an IBV problem.

### 3.2.2 The Heat Equation on a Manifold

Modeling the flow of heat at time  $t$  on a manifold  $\mathcal{M}$ , the heat equation is a second order parabolic partial differential equation [82], and is usually written as

$$\Delta_{\mathcal{M}}U(\mathbf{x}, t) = -\frac{1}{2} \frac{\partial}{\partial t} U(\mathbf{x}, t), \tag{3.13}$$

where  $\Delta_{\mathcal{M}}$  denotes the positive semi-definite Laplace-Beltrami operator of  $\mathcal{M}$ , which is the Riemannian equivalent of the Laplacian (Note: there can be a "conductivity" constant in the Right Hand Side (RHS) that did not used here). The solution  $U(\mathbf{x}, t)$  of the heat equation with initial condition  $U(\mathbf{x}, 0) = u(\mathbf{x})$  describes the amount of heat on the surface at point  $\mathbf{x}$  in time  $t$ .  $U(\mathbf{x}, t)$  is required to satisfy the Dirichlet boundary condition  $U(\mathbf{x}, t) = 0$  for all  $\mathbf{x} \in \partial\mathcal{M}$  and all  $t$ .

The following section, will introduce the Laplace-Beltrami operator and its properties, and the discretization of the Laplace-Beltrami operator, then it will see how Laplace-Beltrami operator will be used to solve the heat equation IBV problem on a manifold.

### 3.3 Laplace-Beltrami

Discrete Laplace-Beltrami operators on triangulated surfaces are a very important ingredient to various applications in geometry processing, including parametrization, modeling, editing, fairing, shape analysis, interpolation, segmentation, remeshing, compression, and matching [83],[84],[27]. Reuter et al. [85] propose to use the set of Laplace-Beltrami eigenvalues spectrum as a shape signature. They show that the spectrum contains enough information to discriminate shapes. Also the Laplace-Beltrami operator is strictly related to the heat diffusion equation, since the discrete heat kernel can be derived from the Laplace-Beltrami eigenfunctions.

#### 3.3.1 Laplace-Beltrami Operator

The Laplace-Beltrami operator is a generalization of the Laplace operator to real-valued (twice differentiable) functions  $f \in \mathcal{C}^2$  on any Riemannian manifold  $\mathcal{M}$  equipped with the Riemannian metric  $g$ , and  $f : \mathcal{M} \rightarrow \mathbb{R}$  be a real-valued function, with  $f \in \mathcal{C}^2$ , defined on a Riemannian manifold  $\mathcal{M}$ . The Laplace-

Beltrami operator can be defined as:

$$\Delta_{\mathcal{M}}f = -div(\nabla f) \quad (3.14)$$

It can be alternatively defined as:

$$\Delta_{\mathcal{M}}f = trace(\nabla^2 f) \quad (3.15)$$

In coordinate notation the Laplace-Beltrami is given by:

$$\Delta_{\mathcal{M}}f = \frac{1}{\sqrt{\det(G)}} \sum_{ij} \frac{\partial}{\partial x_i} f \frac{\partial}{\partial x_j} \det(G) \quad (3.16)$$

where  $G$  is as defined in 3.1.5

If  $\mathcal{M}$  is a domain in the Euclidean plane  $\mathbb{R}^2$ , the Laplace-Beltrami operator reduces to the well-known Laplacian:

$$\Delta_{\mathcal{M}}f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.17)$$

### 3.3.2 Discrete Laplace-Beltrami Operator

The Laplace-Beltrami operator is uniquely defined, whereas the discrete Laplacian can be defined in many ways, and it would not converge to the Laplace-Beltrami operator. In discrete geometry, a distinction is usually made between discrete (or combinatorial) and discretized Laplacian. Discrete Laplacian arises from a discrete object, a triangular mesh representing the shape. Discretized Laplacian, on the other hand, is a consistent numerical approximation of the Laplace-Beltrami operator of the shape, the Laplacian need to be independent or at least minimally dependent on the triangular mesh and thus need the discretized rather than the combinatorial Laplacian. The properties of the continuous Laplace-Beltrami operator will be listed bellow. A consistent discretization



must have properties corresponding to (some subset of) these properties. [86].

**Properties of Smooth Laplacian:**

- (CL1)<sup>1</sup> Constant eigenfunction:  $\Delta_{\mathcal{M}}f = 0$  for any  $f = const.$
- (CL2) Symmetry:  $\langle \Delta_{\mathcal{M}}f, h \rangle_{L_2(\mathcal{M})} = \langle h, \Delta_{\mathcal{M}}f \rangle_{L_2(\mathcal{M})}.$
- (CL3) Locality:  $\Delta_{\mathcal{M}}f$  is independent of  $f(\mathbf{x}')$  for any points  $\Delta_{\mathcal{M}}f \neq \Delta_{\mathcal{M}}f'$  on  $\mathcal{M}$ , or Altering the function value at a distant point will not affect the action of the Laplacian locally.
- (CL4) Euclidean case: if  $\mathcal{M}$  is a part of  $\mathbb{R}^2$ , for any linear function of the form  $f(x, y) = ax + by + c$ ,  $\Delta_{\mathcal{M}}f = 0.$
- (CL5) Maximum principle: harmonic functions (those for which  $\Delta_{\mathcal{M}}f = 0$  in the interior of S) have no local maxima (or minima) at interior points.
- (CL6) Positive semidefiniteness:  $\langle \Delta_{\mathcal{M}}f, h \rangle_{L_2(\mathcal{M})} \geq 0.$

To construct a discrete version of the Laplace-Beltrami operator  $\Delta_{\mathcal{M}}U$ , assume that the shape  $\mathcal{M}$  is sampled at  $N$  points  $\mathbf{x}_1, \dots, \mathbf{x}_N$  and represented as a triangular mesh  $T_{\mathcal{M}}$ . Assuming that a function  $U$  on the shape is discretized and given as a vector with elements  $U_i = U(\mathbf{x}_i)$  for  $i = 1, \dots, N$ . The discrete Laplacian is defined as a linear operator of the form

$$(\Delta_{\mathcal{M}}U)_i = \sum_{j=1}^N w_{ij}(U_i - U_j) \tag{3.18}$$

**Properties of Discrete Laplacian:** To check if the Properties of Smooth Laplacian (CL1)-(CL6) are satisfied for the discrete version in Equation 3.18 test the following properties:

- (DL1)<sup>2</sup> for any  $U = const$ ,  $\Delta_{\mathcal{M}}U = 0.$

---

<sup>1</sup>CL denote continuous Laplacian

<sup>2</sup>DL denote discrete Laplacian

- (DL2) Symmetry:  $\Delta_{\mathcal{M}} = \Delta_{\mathcal{M}}^T$ .
- (DL3) Locality:  $w_{ij} = 0$  if  $i$  and  $j$  do not share an edge.
- (DL4) Euclidean case: if  $\mathcal{M}$  is a part of  $\mathbb{R}^2$ ,  $U(x, y) = ax + by + c$  then

$$(\Delta_{\mathcal{M}}U)_i = \sum_{j=1}^N w_{ij}(U_i - U_j) = 0$$

for all interior vertices (here  $x_j \in \mathbb{R}^2$  denotes the planar coordinates of the  $j^{\text{th}}$  sample).

- (DL5) Positive semidefiniteness:  $\Delta_{\mathcal{M}}U \succeq 0$ .

By satisfying (DL2) and (DL5) it is guaranteed that the eigenvalues of the discrete Laplacians are real and non-negative, and its eigenvectors orthogonal. It is also common to add the following constraint:

- (DL6) Positive weights:  $w_{ij} \geq 0$  for all  $i$ , and for each  $i$ , there exists at least one  $j$  such that  $w_{ij} > 0$ .
- (DL7) Convergence: solution to the discrete PDE involving  $\Delta_{\mathcal{M}}$  converges to the solution of the smooth PDE involving the smooth Laplacians as  $N \rightarrow \infty$  assuming appropriate boundary conditions.

Desbrun et al. [87], proposed a discretization for the Laplace-Beltrami Operator, this discretization is as accurate as the method in [88, 89], so such discretization is a good choice for shape description. According to Desbrun et al. [87], the value of  $\Delta_{\mathcal{M}}U$  at vertex  $\mathbf{x}_i$  is approximated as:

$$(\Delta_{\mathcal{M}}U)_i = -\frac{1}{A_i} \sum_{j \in N_{ei(i)}} (\cot \alpha_{ij} + \cot \beta_{ij})(U_i - U_j), \quad (3.19)$$

where  $(\Delta_{\mathcal{M}}U)_i$  for a mesh function  $U$  denotes its discrete Laplacian evaluated at vertex  $i$  (for  $i = 1, 2, \dots, n$ ),  $n$  is the number of vertices, and  $A_i$  is the

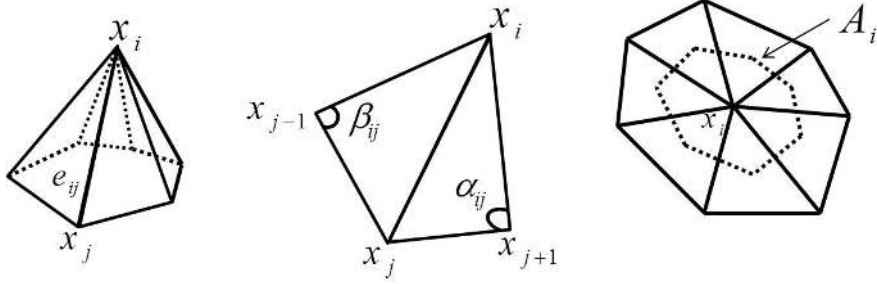


Figure 3.8: Discretization of the Laplace-Beltrami operator using cotangent weights. Left: a vertex  $\mathbf{x}_i$  and its adjacent faces Middle: the definition of the angles  $\alpha_{ij}$  and  $\beta_{ij}$  Right: the definition of the area  $A_i$

Voronoi area at  $i^{\text{th}}$  mesh vertex [90], as shown in Figure 3.8 and  $\alpha_{ij}$ ,  $\beta_{ij}$  are the two angles supporting the edge connecting vertices  $i$  and  $j$ . This discretization preserves many important properties of the continuous Laplace-Beltrami operator, such as positive semi-definiteness, symmetry, and locality, and it is numerically consistent [86].

In a matrix form it can be written as:

$$(\Delta_{\mathcal{M}}U)_i = Lu, \quad (3.20)$$

the mesh Laplace operator  $L$ , is a sparse matrix of size  $n \times n$ , which can be written as  $L = A^{-1}W$  where  $A = \text{diag}(A_i)$ ,  $W = \text{diag}(\sum_{l \neq i} w_{il}) - w_{ij}$ , and  $w_{ij} = (\cot \alpha_{ij} + \cot \beta_{ij})$ . The first  $k$  smallest eigenvalues and eigenfunctions of the Laplace-Beltrami operator discretized according to (3.20) are computed by solving the generalized eigendecomposition problem  $W\phi = \lambda A\phi$ , where  $L = \phi\Lambda\phi^T A$ ,  $\Lambda$  is a diagonal matrix of eigenvalues, and  $\phi$  is a matrix whose columns correspond to the right eigenvectors of  $L$ .

Today, state-of-the-art approaches to shape analysis, synthesis, and correspondence rely on these natural harmonic Laplacian eigenfunctions. The use of Laplacian eigenbases has been shown to be fruitful in many computer graphics applications. Several papers have studied consistent discretizations of the

Laplace-Beltrami operator [86, 91, 92].

An experiment to show the consistency of the Laplace-Beltrami eigenfunctions has been done. The discretization of the Laplace-Beltrami operator was done using the cotangent formula 3.19. The first non-trivial eigenfunctions of the Laplace-Beltrami operator for different deformed shapes from two different classes from the nonrigid world data set [4, 5] are shown in Figure 3.9 and Figure 3.10. This figure shows that the first non-trivial eigenfunctions are invariant to the shape deformation.

Figure 3.11 and Figure 3.12 show the first four and the first three eigenfunctions of the Laplace-Beltrami operator of human shapes with different deformations. The figure clearly demonstrates that the eigenfunctions are invariant to the shape deformation, and that they can be even used for mesh segmentation [6].

Figure 3.13 and Figure 3.14 show the spectral projection [6] of the first three eigenfunctions of the Laplace-Beltrami of human and dog shapes under different deformation. Notice the difference in the projection of the human shape and for the dog shape. The figure clearly demonstrates that the eigenfunction's spectral projection are invariant to the shape deformation, and that can be even used as a shape model for recognition or shape normalization.

However, many applications involving multiple shapes are obstructed by the fact that Laplacian eigenbases computed independently on different shapes are often incompatible with each other [93]. It requires an approximate joint diagonalization for the construction of coupled bases of the Laplacians of multiple shapes.

### 3.4 Heat Operator and Heat Kernel

This section, introduces the basics of diffusion on Riemannian manifolds that are necessary to define the purposed heat kernel signature. The shape will be modeled as a Riemannian manifold possibly with boundary. Two methods

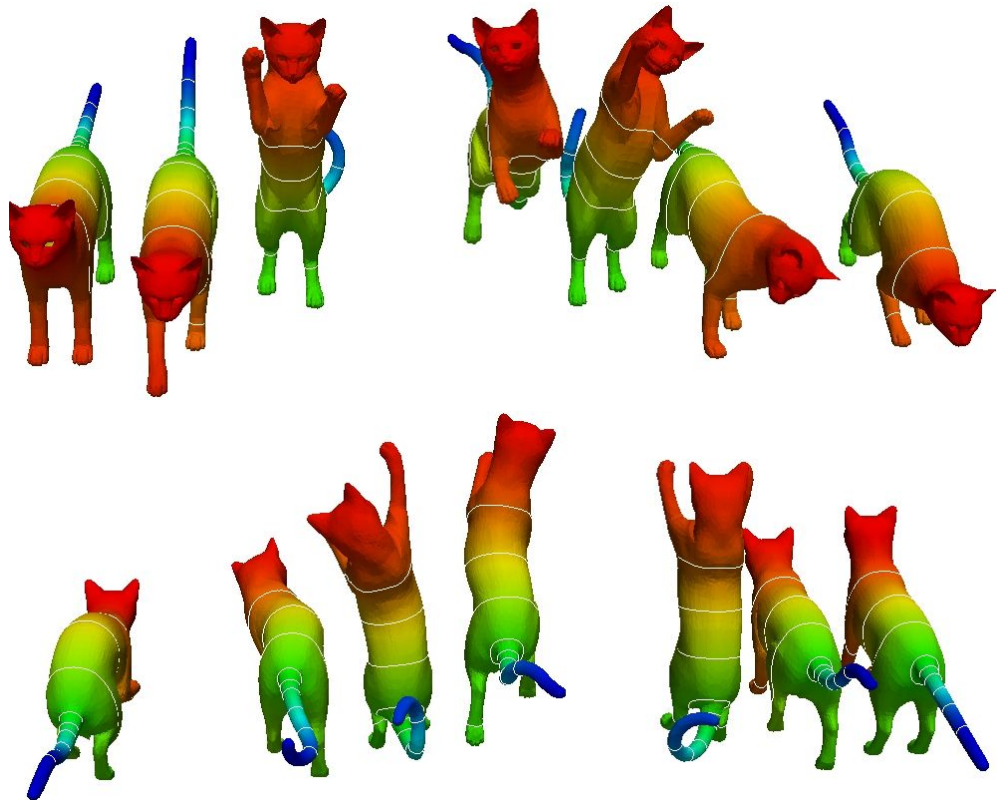


Figure 3.9: The first non-zero eigenfunctions of the Laplace-Beltrami operator of cat shape from the nonrigid world data set [4, 5] with two different views. Colors represent the values of the eigenfunction at each point of the shape. This illustration shows that the first non-trivial eigenfunctions are invariant to the shape deformation.

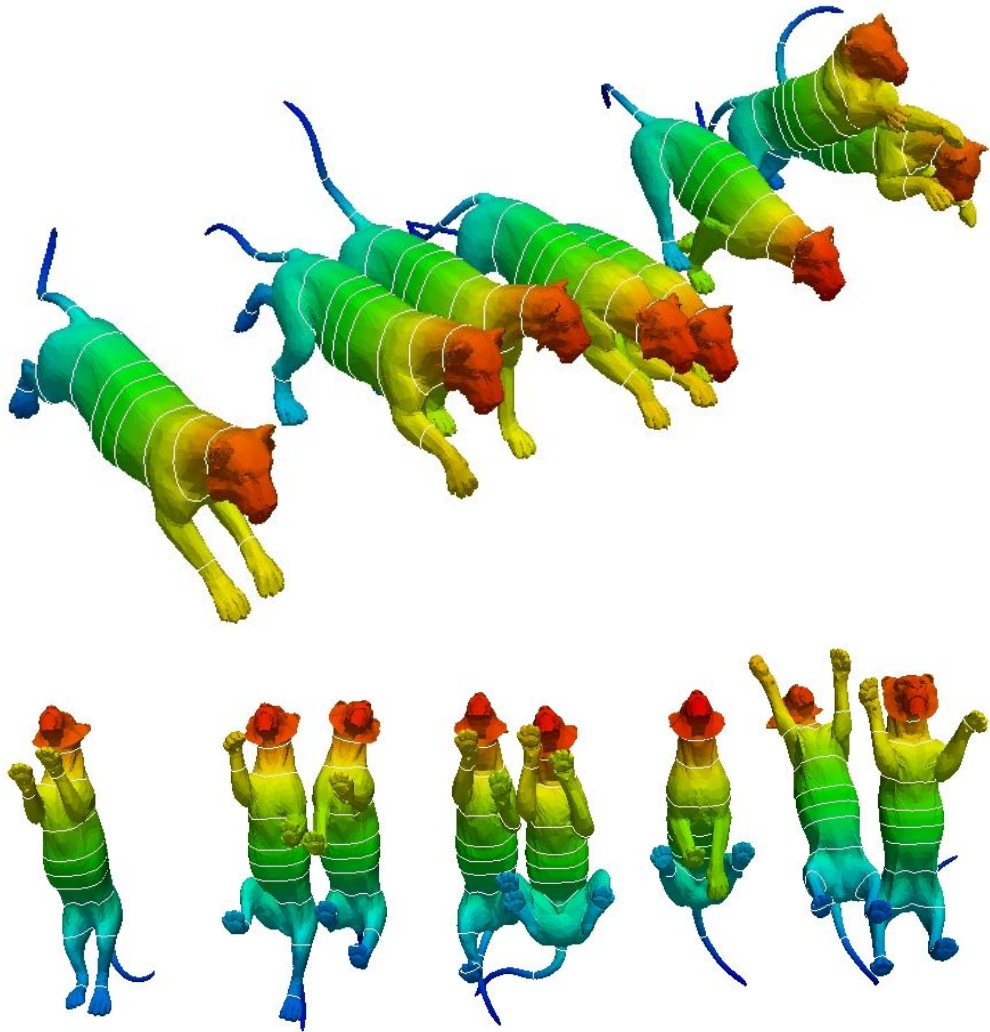


Figure 3.10: The first non-zero eigenfunctions of the Laplace-Beltrami operator of lion shape with two different views. Colors represent the values of the eigenfunction at each point of the shape. This illustration shows that the first non-trivial eigenfunctions are invariant to the shape deformation.

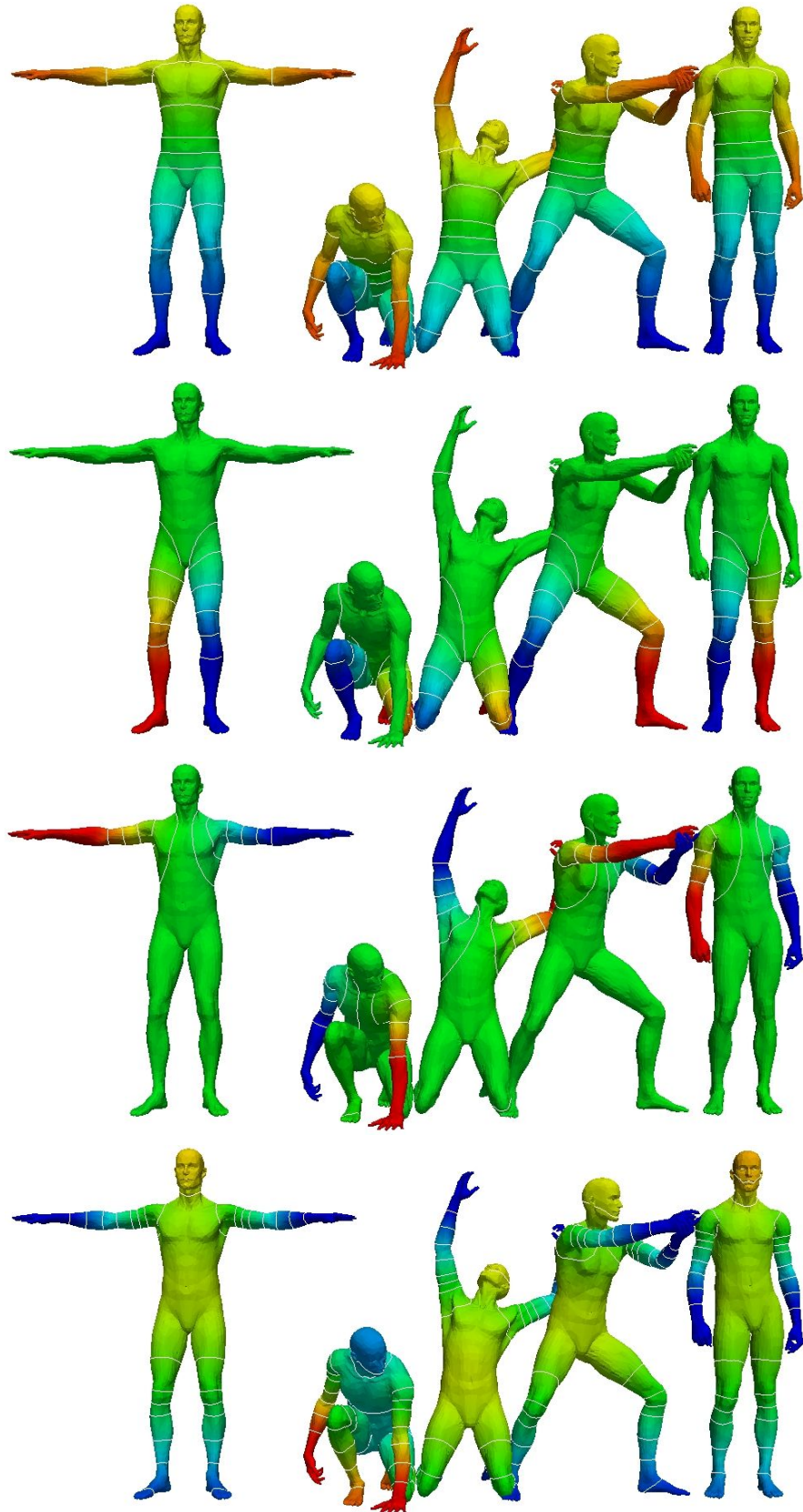


Figure 3.11: The first four eigenfunctions of the Laplace-Beltrami operator of human shapes with different deformation. This illustration shows that the eigenfunctions are invariant to the shape deformation and can be used for shape segmentation as [6].

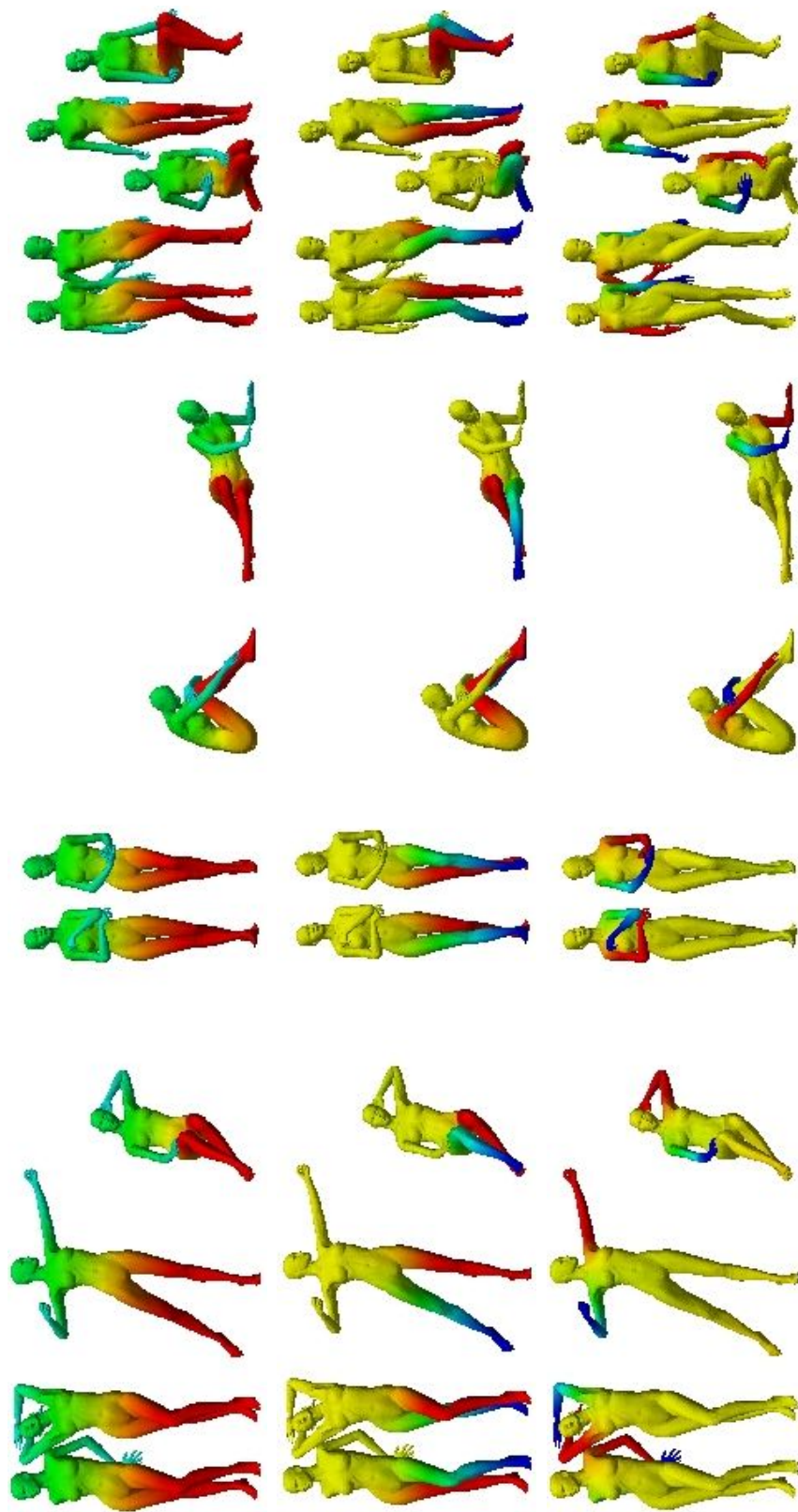


Figure 3.12: The first three eigenfunctions of the Laplace-Beltrami operator of women shapes with different deformation. This illustration shows that the eigenfunctions are invariant to the shape deformation and can be used for shape segmentation as [6].



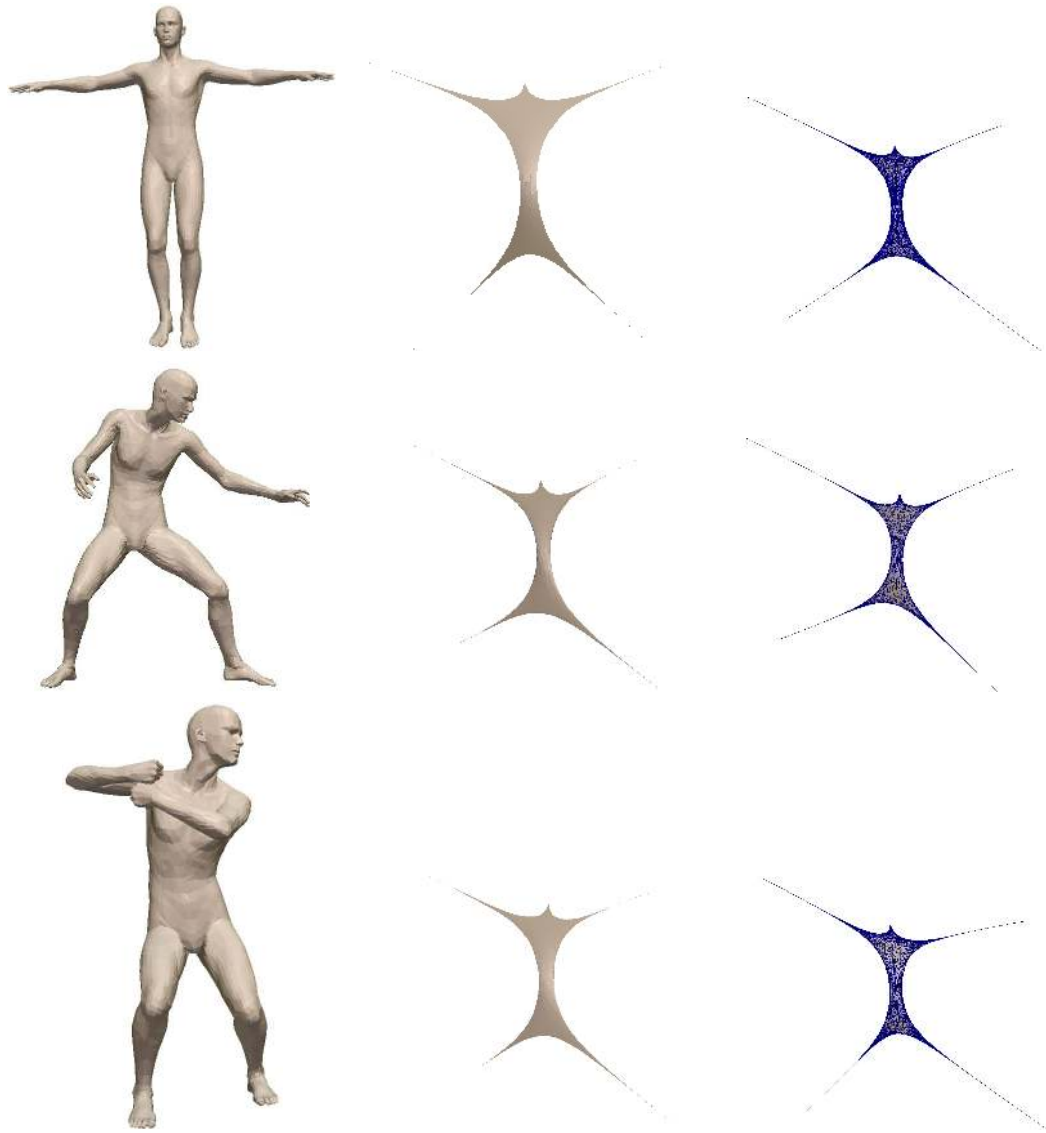


Figure 3.13: Spectral projection onto the first 3 eigenfunctions of the Laplace-Beltrami of human shape with different deformation. Left: the 3D shape, Middle: the spectral projection as a surface, Right: the spectral projection as a triangulated mesh.

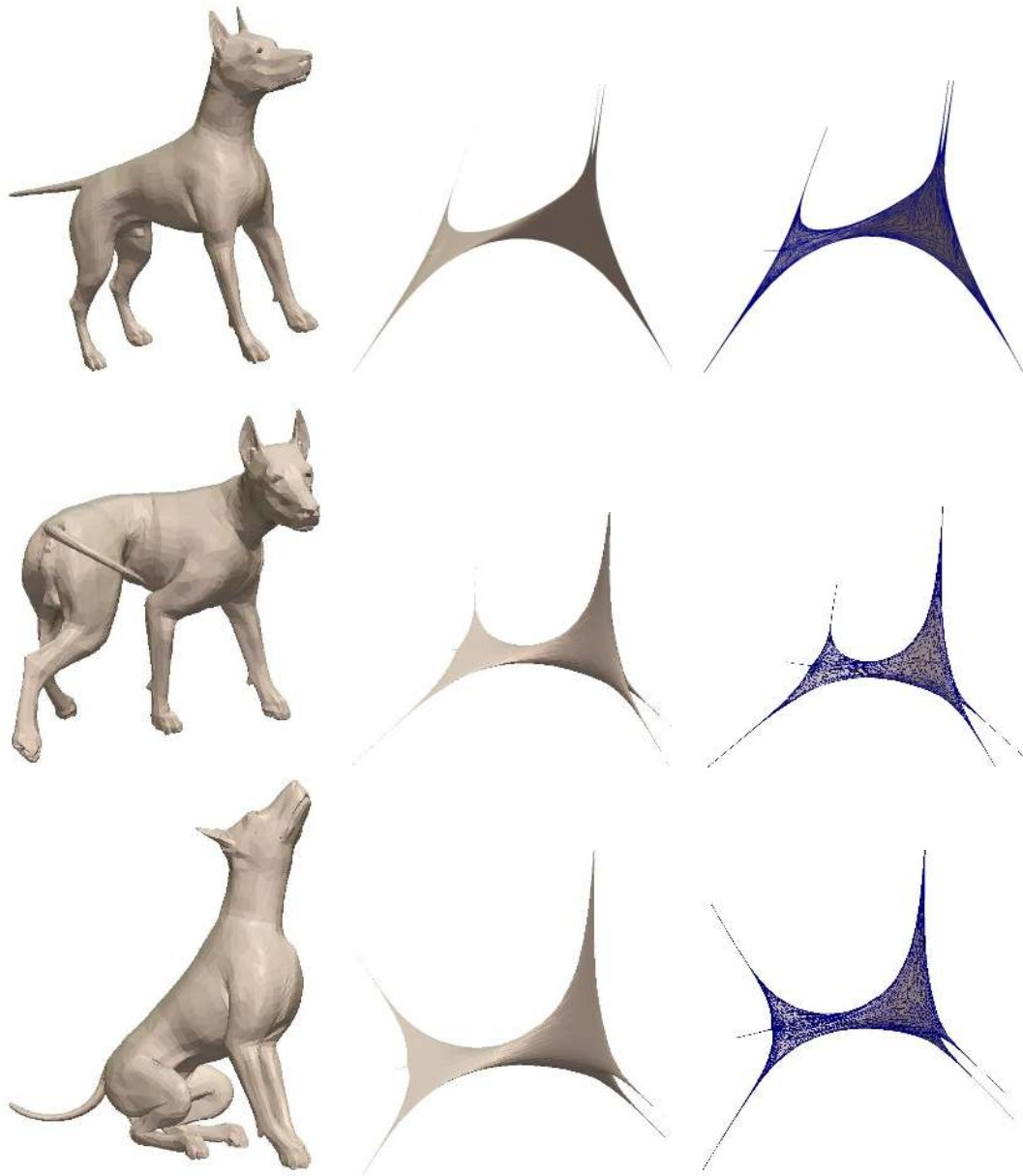


Figure 3.14: Spectral projection onto the first 3 eigenfunctions of the Laplace-Beltrami of dog shape with different deformation. Notice the difference in the projection over this of the human shape in Figure 3.13 Left: the 3D shape, Middle: the spectral projection as a surface, Right: the spectral projection as a triangulated mesh.

will be shown to reconstruct the heat kernels: the first method is based on the property that Laplace- Beltrami operator and the heat operator share the same eigenfunctions, and the second method is based on linear FEM approximation of the heat equation to derive a discrete heat kernel.

### 3.4.1 Heat Kernel

The heat diffusion propagation over  $\mathcal{M}$  is governed by the heat equation (3.13). Given an initial heat distribution  $u : \mathcal{M} \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ , as a scalar function defined on a compact manifold  $\mathcal{M}$ , the scale based representation  $U : \mathcal{M} \times \mathbb{R} \rightarrow \mathbb{R}$  of  $u$  with,  $\lim_{t \rightarrow 0} U(\mathbf{x}, t) = u(\mathbf{x})$  satisfies the heat equation for all  $t$ . To interpret the heat operator and the heat kernel in the discrete setting, the heat equation on a mesh that approximates the underlying manifold can be re-written as:

$$\frac{\partial}{\partial t} U(\mathbf{x}, t) = e^{-tL} u(\mathbf{x}) \quad (3.21)$$

When  $L = \phi \Lambda \phi^T A$  as stated before, then  $e^{-tL} = \phi e^{-t\Lambda} \phi^T A$

The heat kernel can be thought of as the amount of heat that is transferred from  $\mathbf{x}$  to  $\mathbf{y}$  in time  $t$  given a unit heat source at  $\mathbf{x}$ . Since  $\mathcal{M}$  is compact then the solution  $U(\mathbf{x}, t)$  can be written as  $U(\mathbf{x}, t) = \int_{\mathcal{M}} K(\mathbf{x}, \mathbf{y}, t) u(\mathbf{y}) dy$ . where  $K(\mathbf{x}, \mathbf{y}, t)$  is the heat kernel.

Then the heat kernel  $K(\mathbf{x}, \mathbf{y}, t) = e^{-tL}$  has the following spectral decomposition

$$K(\mathbf{x}, \mathbf{y}, t) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \phi_i(\mathbf{x}) \phi_i(\mathbf{y}), \quad (3.22)$$

where  $\lambda_i$  and  $\phi_i$  are the  $i^{th}$  eigenvalue and the  $i^{th}$  eigenfunction of the Laplace-Beltrami operator respectively, and  $\mathbf{x}$  and  $\mathbf{y}$  denote two vertices. The heat kernel 3.22 is one way to approximate the heat operator based on the spectral de-

composition of the Laplace-Beltrami operator. In the following, a more accurate method for the discretization of the heat kernel based on the weighted linear Finite Elements Method (FEM) [94] will be described. FEM gives a way of discretizing (i.e. finding a discrete approximation) of a continuous equation.

### 3.4.2 Finite Element Discretization of the Heat Equation

The weak formulation of the heat equation (3.13) through the Galerkin formulation with test function  $\varphi \in \mathcal{C}^2$  yields

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U(\mathbf{x}, t) d\sigma + \frac{1}{2} \int_{\mathcal{M}} \varphi \Delta_{\mathcal{M}} U(\mathbf{x}, t) d\sigma = 0, \quad (3.23)$$

where  $d\sigma$  is the Riemannian measure, and the integration is carried out over the manifold  $\mathcal{M}$ . Then, using the Green formula Equation 3.23 will be:

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U d\sigma + \frac{1}{2} \int_{\mathcal{M}} \operatorname{div}(\nabla U) \varphi d\sigma = 0, \quad (3.24)$$

or equivalently,

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U d\sigma + \frac{1}{2} \int_{\mathcal{M}} (\nabla U) \cdot \nabla \varphi d\sigma = 0. \quad (3.25)$$

Thus the weak formulation of (3.13) can be written as:

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U d\sigma + \frac{1}{2} \int_{\mathcal{M}} (\nabla U \cdot \nabla \varphi) d\sigma = 0 \quad (3.26)$$

Let  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$  be a family of  $n$  linearly independent  $C^2$  functions. An approximation  $\tilde{U}(\mathbf{x}, t)$  to  $U(\mathbf{x}, t)$  can be defined as:

$$\tilde{U}(\cdot, t) := \sum_{i=1}^n a_i(t) \varphi_i, \quad t > 0. \quad (3.27)$$

Replacing  $U$  by  $\tilde{U}$  in (5.6) then:

$$\sum_{i=1}^n \frac{\partial}{\partial t} a_i(t) \int_{\mathcal{M}} \varphi_i \varphi_j d\sigma + \frac{1}{2} \sum_{i=1}^n a_i(t) \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) d\sigma = 0, \quad (3.28)$$

or simply as:

$$\sum_{i=1}^n B(i, j) \frac{\partial}{\partial t} a_i(t) + \frac{1}{2} \sum_{i=1}^n L(i, j) a_i(t) = 0 \quad (3.29)$$

where

$$B(i, j) = \int_{\mathcal{M}} \varphi_i \varphi_j d\sigma, \text{ and} \quad (3.30)$$

$$L(i, j) = \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) d\sigma. \quad (3.31)$$

To discretize Equation 3.29 let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Assume linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  that are equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolate between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $t_k$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at vertices  $i, j, k$  respectively, and  $|t_k|$  is the area of  $t_k$ , then it can be shown that:

$$\begin{aligned} \nabla \varphi_i \cdot \nabla \varphi_j &= -\frac{\cot \alpha_k}{2 |t_k|} \\ |\nabla \varphi|^2 &= \frac{\cot \alpha_j + \cot \alpha_k}{2 |t_k|} \\ \int_T \varphi_i d\sigma &= \frac{|t_k|}{3} \\ \int_T \varphi_i^2 d\sigma &= \frac{|t_k|}{6} \\ \int_T \varphi_i \varphi_j d\sigma &= \frac{|t_k|}{12} \end{aligned} \quad (3.32)$$

For the proofs of ( 3.32) see Appendix B.

Then  $B$  will be:

$$B(i, j) = \begin{cases} \frac{|t_r|+|t_s|}{12} & \text{if } j \in N_e(i) \\ \frac{\sum_{k \in N_e(i)} |t_k|}{6} & \text{if } i = j \end{cases} \quad (3.33)$$

$L(i, j)$  is given by

$$L(i, j) = \begin{cases} w(i, j) = \frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2} & \text{if } j \in N_e(i) \\ -\sum_{k \in N_e(i)} w(i, k) & \text{if } i = j \\ 0 & \text{else} \end{cases} \quad (3.34)$$

where  $N_e(i)$  is the 1-star neighbors of vertex  $i$ ,  $|t_k|$  is the area of triangle  $t_k$ , and  $t_r$  and  $t_s$  are the triangles that share the edge  $(i, j)$ . For the proofs of 3.33 and 3.34 see Appendix C.

To compute the solution to (3.13), consider the generalized eigensystem  $\{\lambda_i, \phi_i\}_{i=1}^n$  of  $(L, B)$ , which satisfies the relations  $L\phi_i = \lambda_i B\phi_i, i = 1, \dots, n$ . Since the Laplacian eigenvectors  $\{\phi_i\}_{i=1}^n$  form a basis of  $\mathbb{R}^n$  and  $(\tilde{U}(\mathbf{x}, t))_{i=1}^n \in \mathbb{R}^n$ , for any  $t \in \mathbb{R}^+$  the solution  $\tilde{U}(\mathbf{x}, t)$  can be expressed as  $\tilde{U}(\cdot, t) := \sum_{i=1}^n a_i(t)\phi_i$  where  $\mathbf{a} = (a_i(t))_{i=1}^n$  is the unknown vector.

After solving for the coefficients  $a_i(t)$ , then  $a_i(t) = \exp(-\frac{1}{2}\lambda_i t)\langle u(\mathbf{x}), \phi_i \rangle_B$  where  $u(\mathbf{x})$  is the initial value of  $\tilde{U}(\mathbf{x}, t)$ . Then:

$$\tilde{U}(\cdot, t) := \sum_{i=1}^n \exp(-\frac{1}{2}\lambda_i t)\langle u(\mathbf{x}\phi_i) \rangle_B \phi_i, \quad (3.35)$$

or in a matrix form:

$$\tilde{U}(\cdot, t) := \phi D(t) \phi^T B u(\mathbf{x}) \quad (3.36)$$

where  $\phi = [\phi_1, \phi_2, \dots, \phi_n]$ , and

$$D(t) = \text{diag}(\exp(-\frac{1}{2}\lambda_1 t), \exp(-\frac{1}{2}\lambda_2 t), \dots, \exp(-\frac{1}{2}\lambda_n t)).$$

Then the heat kernel will be:

$$K(\mathbf{x}, \mathbf{y}, t) := \phi D(t) \phi^T B, \quad (3.37)$$

### 3.4.3 Properties of the Heat Kernel

The heat kernel  $K(\mathbf{x}, \mathbf{y}, t)$  has many good properties [74, 95]:

- $K(\mathbf{x}, \mathbf{y}, t)$  is symmetric.
- It is invariant under isometric deformations, which is a direct consequence of the invariance of the Laplace-Beltrami operator.
- It is informative: by only considering its restriction to the temporal domain, a concise and informative signature can be obtained .
- It is multi-scale: for different values of  $t$  the heat kernel reflects local properties of the shape around  $x$  at small  $t$  and the global structure of  $M$  from the point of view of  $x$  at large values of  $t$ .
- And it is stable under perturbations of the underlying manifold.

[74] proved that heat kernel is closely related to curvature, diffusion distance [73], and the Global Point Signature (GPS) [28].

### 3.4.4 Heat Kernel Signature

Sun et al. [74] proposed to use the HKS as a local shape descriptor, and it can be written as:

$$K(\mathbf{x}, t) = K(\mathbf{x}, \mathbf{x}, t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \phi_k(\mathbf{x})^2, \quad (3.38)$$

The HKS has several desired properties [74]: it is intrinsic and thus isometry-invariant (two isometric shapes have equal HKS), multi-scale and thus captures

both local features and global shape structure, and also informative: under mild conditions, if two shapes have equal heat kernel signatures, they are isometric.

### 3.4.5 Temperature Distribution Descriptor (TD)

In [76], the heat mean signature (HMS) was defined for quantitatively evaluating the temperature distribution resulting from the heat flow process. The HMS has the following formulation:

$$HMS(i) = \frac{1}{N} \sum_{j, j \neq i} H_t(i, j), \quad (3.39)$$

where  $i$  and  $j$  denote the  $i^{th}$  and  $j^{th}$  vertex on the surface, respectively and  $N$  is the total number of vertices.  $HMS(i)$  can be physically interpreted as the average temperature on the surface obtained by applying a unit amount of heat on the vertex  $i$  and after a certain amount of time of heat dissipation. TD descriptor is constructed based on the distribution of the values of heat mean signature over the meshed surface as in [1]. The TD descriptor will be used for the sake of comparison with the proposed method in the next chapter.

## 3.5 Summary

This chapter gave the derivation of the heat equation from the physical properties of the object, and from the rate of heat flow through the boundaries of the object. Then it described the Initial Boundary Value Problem (IBV). Also this chapter gave the details of the Laplace-Beltrami operator, its properties, the discretization of the Laplace-Beltrami operator, and how to use it to solve the heat equation IBV problem on a manifold. After that it talked about the solution of the heat equation and the basics of diffusion on Riemannian manifolds that are necessary to define the heat kernel. Finally it described the heat kernel signature and its properties.



## CHAPTER 4

### PROPOSED SHAPE DESCRIPTORS

Object matching algorithms consist of three major parts: feature detection, feature description, and a feature matching. Figure 4.1 illustrates a typical approach of object recognition using local features.

Several interest point detectors have been developed in the literature (see Section 2.2.1). For every detected interest point, a numerical vector called the "feature descriptor" is built to describe that point, or a patch (local neighborhood), around it.

To be able to correctly check the similarity of non-rigid shapes, it is needed to look for the properties that distinguish between these shapes. Such properties are called deformation invariant signatures, and the similarity measure based on these properties is called deformation-invariant similarity. The shape signature (or descriptor) captures the properties of the shape that distinguish it from shapes belonging to other classes. This chapter looks for deformation invariant signatures and deformation-invariant similarity measure to model non-rigid 3D objects.

This chapter proposes an approach for shape modeling using heat kernels (HK). This shape model will be used for shape matching and retrieval. It proposes two different types of shape descriptors:



Figure 4.1: The three parts of object matching algorithms.

The first one is based on the histograms of the HK [30]. The HK calculated at a small a number of critical points detected on the shape surface. The HK is calculated at different time scales, then is combined with the *normalized* eigenvalues of the Laplace-Beltrami operator. The proposed descriptors are then used for classification via the collaborative representation-based classification with regularized least square (CRC-RLS) algorithm.

The second descriptor is based on scale-invariant heat kernels [31]. The first non-trivial Laplace-Beltrami eigenfunction is used to detect a small number of sparse critical points on the shape surface. These points are robust to the shape class, and their number can in itself be used as one of the discriminatory features among the various classes. The HK is calculated only at the detected critical points at different time scales. This reduces the size or dimensionality of the descriptors, yet does not sacrifice the accuracy of the presentation. Then scale invariance is achieved using a novel transformation method that is simpler and considerably less noise sensitive than the method proposed in [3]. A concatenation of the histograms of the significant components of the scale-invariant HK for all the points, combined with the *normalized* eigenvalues of the Laplace-Beltrami operator, is used as a feature vector for classification. The resulting descriptor is then used for classification via the relaxed collaborative representation (RCR) model [96, 97]. The resulting descriptor captures the local as well as global shape information since it uses the temperature distribution at the critical points at several time samples.

The proposed descriptors are compact in size, and efficient in computation. For comparison the proposed approaches are compared to state-of-the-art approaches on two different data sets: the nonrigid world data set and the SHREC 2011. The results have indeed confirmed the improved performance of the sparse representation over the original non-reduced descriptors, yet reducing the time and space complicity for the shape retrieval problem.

The following section describes the proposed method for 3D feature Points Detection. Then it gives the details of the two shape descriptors.

## 4.1 Critical Points Detection

For a piecewise linear real-valued function  $\phi$  given by the values at the vertices ( $\phi_i$ ) of a triangle mesh, a critical point is defined as a vertex  $i$  whose function  $\phi_i$  is a maximum or minimum over its neighborhood (in two rings). These points are detected using the local maxima/minima of the first non-trivial Laplace-Beltrami eigenfunction [6]. Critical points detected near the boundary are discarded. Figure 4.2 shows the critical points detected from the first non-trivial eigenfunction for several sample shapes. The figure gives the total number of critical points for each shape. It is interesting to observe that shapes belonging to the same class consistently have almost the same number of critical points, whereas these numbers differ from one class to another. As such, this number can be used as one of the discriminatory features between the different classes.

## 4.2 Shape Descriptor Based on Heat Kernels

This section presents an approach for shape matching and retrieval based on scale-invariant heat kernel (HK). Several aspects are novel in this approach. The problem of scale invariance found in HKS as proposed in [74] is solved by normalizing the Laplace-Beltrami eigenvalues and eigenfunctions. This section also proposes a new shape descriptor. The first non-trivial Laplace-Beltrami eigenfunction is used to detect a small number of sparse critical points on the surface of the shape. These points are robust to the shape class, and their number can in itself be used as one of the discriminatory features among the various classes. The shape descriptor consists of the HK at the detected critical points for different time samples, combined with the *normalized* eigenvalues of the Laplace-Beltrami

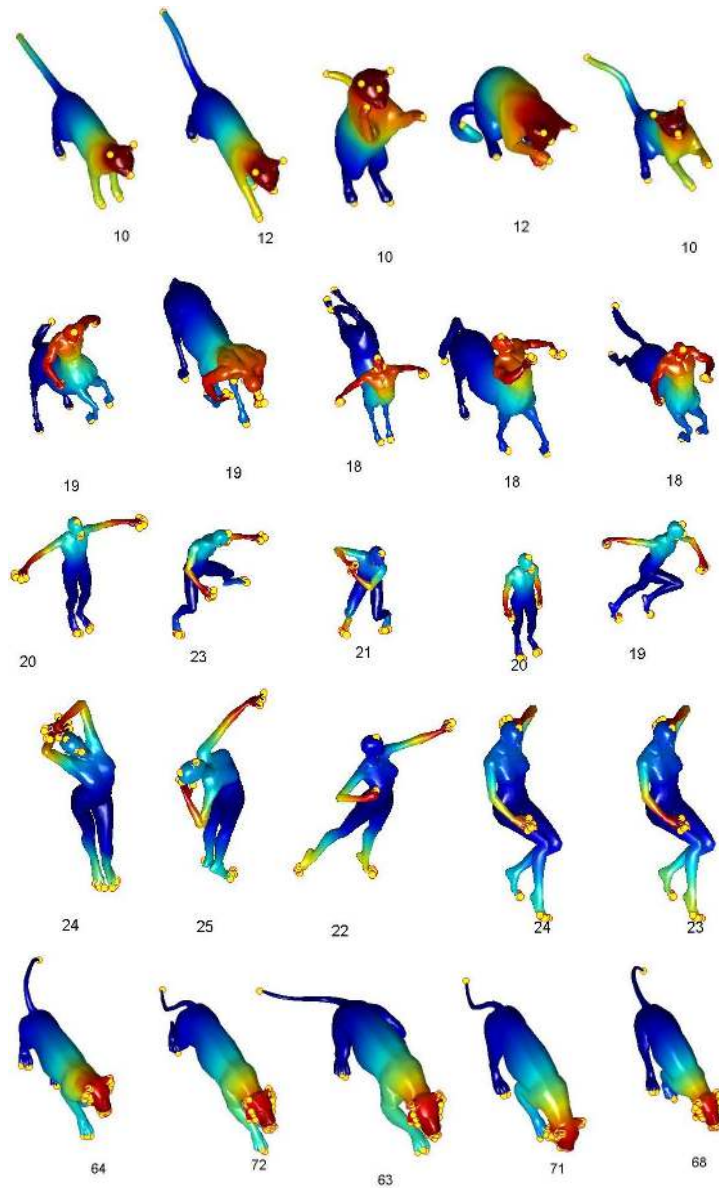


Figure 4.2: Critical points detected for different shapes with different deformation. Number below each shape represents the total number of shape critical points. Colors indicate the average temperature induced from these critical points throughout all shape vertices.

operator. Sparse representation [97] is used to reduce the descriptor dimensionality. The resulting descriptor is then used for classification via the collaborative representation-based classification with the regularized least square (CRC-RLS) algorithm [97]. The proposed descriptor is compact in size (the feature vector can be down to 64), and efficient in computation (since it is constructed from the HK at sparse points). This descriptor increases the accuracy of the recall, yet reduces the time and space complexity for the shape retrieval problem. It captures the local as well as global shape information since it uses the HK at several time scales. Further go to the extreme it will be shown that this descriptor computed at just one critical point still works rather well for shape retrieval. The proposed approach is compared to the shape-DNA [2] and the TD approach of [1] on two different 3D data sets, the nonrigid world data set [4, 5] and the SHREC 2011-Shape Retrieval Contest of Non-rigid 3D Watertight Meshes [98]. This approach achieves state-of-the-art results on the two data sets.

The most related work to this approach is the recent Shape Google approach [3]. However ours is different in several aspects: 1) Shape Google uses dense or sparse descriptors. The sparse descriptor is closer to this work, but requires the calculation of HKS on all vertices to detect critical points. However in the proposed approach critical points are detected from the first non-trivial Eigenfunctions of the Laplace-Beltrami operator which is already calculated, this saves time. 2) Scale-invariance of HKS is achieved by using Fourier magnitude to avoid scale detection. In the current case it is achieved by normalizing Laplace-Beltrami eigenvalues and Eigenfunctions. 3) Shape Google uses the Bag of Features (BoF), but this method uses the CRC-RLS algorithm [97]. Overall, the proposed sparse representation is more compact, efficient to compute, and still discriminative.

The HK construct is constructed at a given vertex  $\mathbf{x}$  using:

$$K(\mathbf{x}, \mathbf{y}, t) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \phi_i(\mathbf{x}) \phi_i(\mathbf{y}), \quad (4.1)$$

where  $\lambda_i$  and  $\phi_i$  are the  $i^{th}$  eigenvalue and the  $i^{th}$  eigenfunction of the Laplace-Beltrami operator respectively, and  $\mathbf{x}$  and  $\mathbf{y}$  denote two vertices. As described in Chapter 3 Section 3.4.1.

Figure 4.3 shows the HK induced from one critical point for one sample shape at different time samples. For a smaller amount of time, the signature captures local shape information. As time elapses, the signature tends to capture more global shape details.

### 4.2.1 The Heat Kernel Descriptor

Figure 4.4 shows the steps for the construction of the proposed descriptor. The proposed descriptor of a given shape is constructed as follows: HKs are calculated for all critical points at various (about 100) time samples. Since complexity of using the heat kernel as a signature is extremely high, and it would be difficult to compare signatures of two different points, the histograms and sparse representation is used to overcome signature alignment problem and reduce the descriptor size. At each time sample, a histogram of 100 pin is calculated, then all the histograms for different times and from all detected critical points are concatenated to build a long feature vector. Then the normalized eigenvalues of the Laplace-Beltrami operator (the nonzero diagonal elements of  $\Lambda$ ) are appended to this vector. This vector, dubbed Critical Points-based Temperature Distribution (CP-TD), can be used for shape retrieval in its current form, or reduced to a sparse representation of a lower dimensionality, then used for classification using the collaborative representation-based classification with regularized least square (CRC-RLS) algorithm [97], which is described in the next section. This sparse descriptor is called Sparse Representation of CP-TD (SRCP-TD).

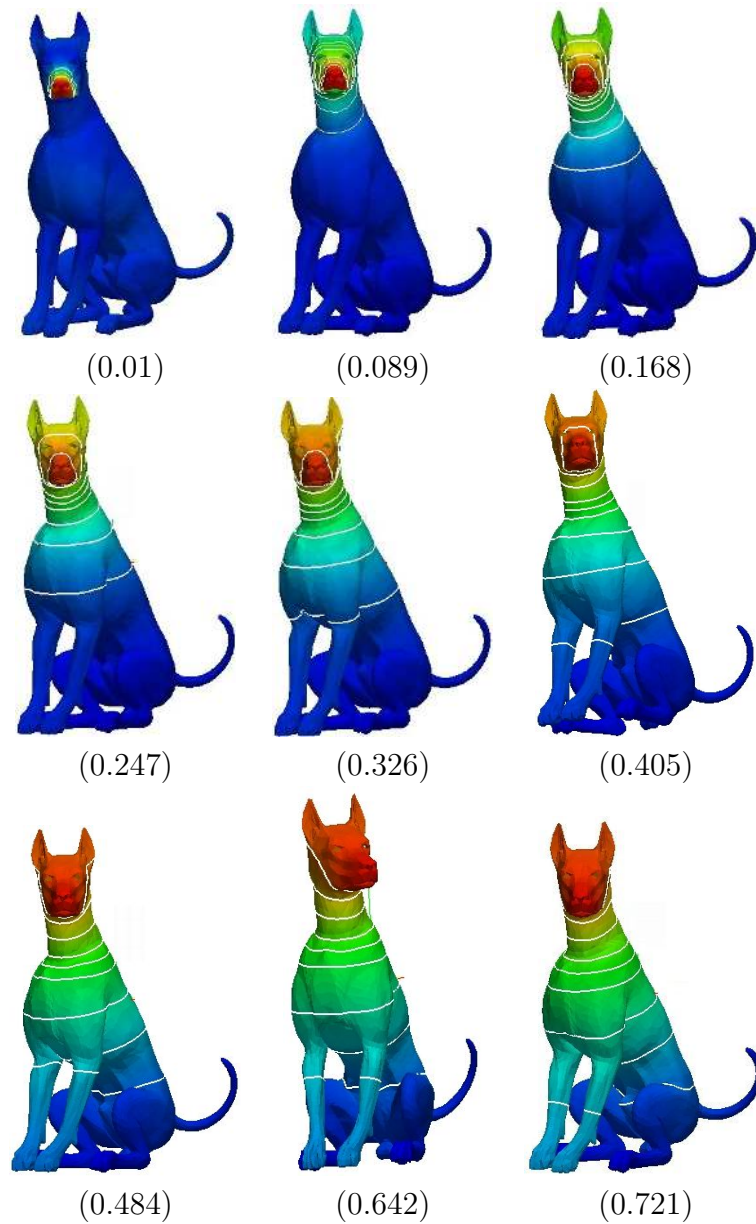


Figure 4.3: HKS induced from one critical point detected at the nose of the dog shape at different time samples. The shown numbers indicate the time samples  $t$ . Colors represent the values of the heat induced from the source point to each vertex of the shape. For a smaller amount of time, the signature is more local. As time progresses, the signature becomes more global.

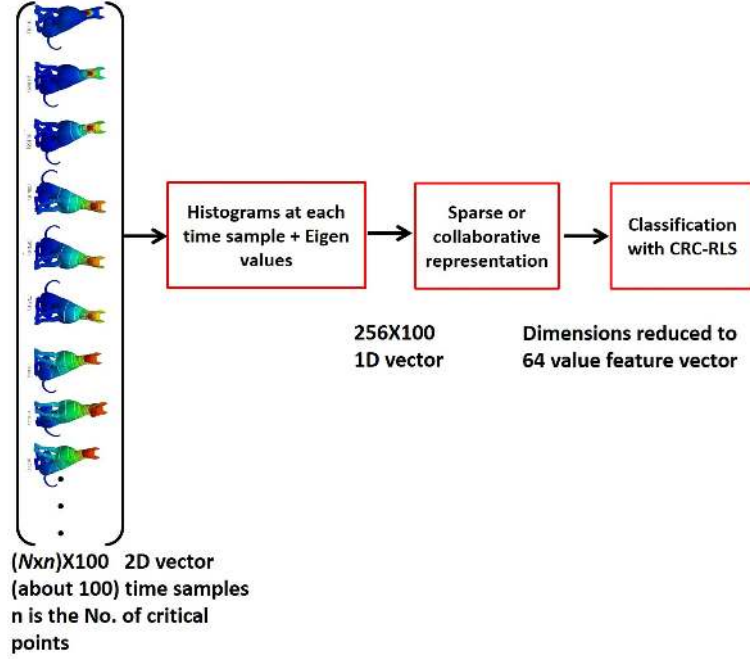


Figure 4.4: Construction of the proposed descriptor (N is the number of object vertices).

### 4.2.2 Sparse Representation of HK

As a recently proposed technique, sparse representation based classification (SRC) has been widely used for face recognition (FR). Recently, sparse representation has also been used in pattern classification. Huang et al. [99] sparsely coded a signal over a set of redundant bases and classified the signal based on its coding vector.

Sparse representation (or coding) codes a signal  $y$  over a dictionary  $\phi$  such that  $y = \phi\alpha$  and  $\alpha$  is a sparse vector. The sparsity of  $\alpha$  can be measured by  $l_0$ -norm, which counts the number of non-zeros in  $\alpha$ . Since the combinatorial  $l_0$ -minimization is NP-hard, the  $l_1$ -minimization, as the closest convex function to  $l_0$ -minimization, is widely employed in sparse coding:  $\min_{\alpha} \|\alpha\|_1 \text{ s.t. } \|y - \phi\alpha\|_2 < \varepsilon$ , where  $\varepsilon$  is a small constant. Zhang [97] indicated that it is the CR, but not the  $l_1$ -norm sparsity, that plays the essential role for classification in SRC, and proposed a new classification scheme, namely collaborative representation



(CR) based classification with regularized least square (CRC-RLS), which has significantly less complexity than SRC but leads to very competitive classification results. The same approach is used to reduce the dimensionality of the proposed descriptor and improve the classification results.

### 4.2.3 Experimental Results

To test the performance of the proposed descriptors two data sets are used. The first one is the nonrigid world data set [4, 5]. The database contains a total of 148 3D nonrigid objects in a variety of poses, including 9 cats, 11 dogs, 3 wolves, 17 horses, 15 lions, 21 gorillas, 1 shark, 24 female figures, and two different male figures, containing 15 and 20 poses. The database also contains 6 centaurs, and 6 seahorses for partial similarity experiments. Each object contains approximately 3400 vertices. The second database is the SHREC 2011 - Shape Retrieval data set [98]. This is a large-scale database consists of 600 non-rigid 3D objects that are derived from 30 original models.

For the sake of comparison, this work shows the results of the Shape-DNA approach [2], describing shapes by the vector of the first eigenvalues of the Laplace-Beltrami operator. The first 15 eigenvalues are used to construct the Shape-DNA descriptors. Eigenvalues were computed using the same cotangent weight discretization. Also the proposed approach is compared to the method in [1] that uses the (TD) as a shape descriptor.

Figure 4.5 shows sample shape retrieval results of the SRCP-TD descriptor on the nonrigid world data set using the CRC-RLS algorithm [97]. The retrieved objects per each query are ranked from left to right based on the distance measure of the classification algorithm. The figure clearly demonstrates the high retrieval rate of the approach. It is interesting to note that the objects similar to the query pose are retrieved first, then come objects of the same shape but with different poses. This shows that this descriptor can be used for object pose estimation as

well. The wrong matches in the first, second and last row (colored in green) are only retrieved after all objects of correct shape have been already retrieved.

Figure 4.6 shows sample shape retrieval results of the same SRCP-TD descriptor on the SHREC 2011 data set using the CRC-RLS algorithm as well. The figure shows the first 15 matches for each query ranked according to the distance measure of the CRC-RLS algorithm. Afterwards, the objects ranked from 30 – 35 for each query are shown on the right. Several of these objects are also similar in shape to the query object.

For the sake of quantitative assessment of the approach performance, the following standard five evaluation measures are recorded (see [100] for detailed definitions): Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-measure (E), and Discounted Cumulative Gain (DCG). Table 4.1 lists these measures on the nonrigid world data set. The table compares the proposed descriptor (SRCP-TD) against the TD and the shape-DNA descriptors using two different classifiers: the K-NN classifier and the CRC-RLS algorithm. The SRCP-TD using the CRC-RLS algorithm consistently provides the highest scores. Analogously, Table 4.2 shows the performance on the SHREC 2011 data set. The SRCP-TD descriptor again exhibits the best performance but using the K-NN classifier this time.

In the previous experiments, all the critical points detected on a shape have been used to construct the shape descriptor. Table 4.3 gives the five evaluation measures on the non-rigid world data set computed for the proposed descriptor at only one critical point. This point is located at the minimum of the heat mean signature of the whole shape. Figure 4.7 shows the detected single point on sample shapes of the non-rigid world data set. However the measures still reflect the high performance of the proposed descriptors. Comparing this results to Table 4.1 it is clear that the descriptor at only one critical point achieves better performance for the NN measure. But detecting this single point requires to calculate the heat mean signature of the whole shape which is computationally complex.

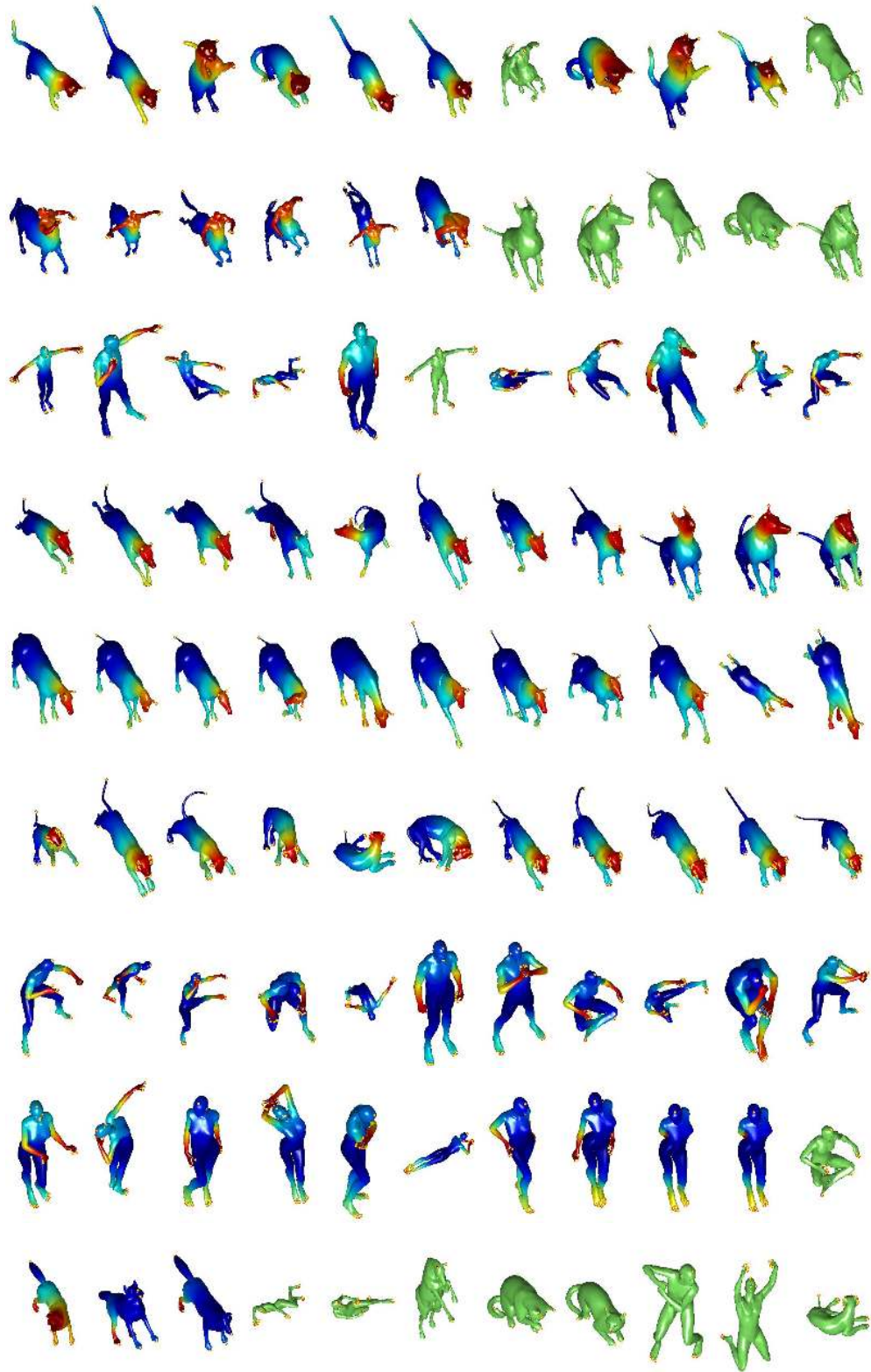


Figure 4.5: Some shape retrieval results on the nonrigid world data set. Left: Queries. Right: First 10 matches using the SRCP-TD descriptor. The color represents the mean temperature induced from only the detected critical points. The detected sparse points are shown in yellow. The green-colored shapes are wrong matches.

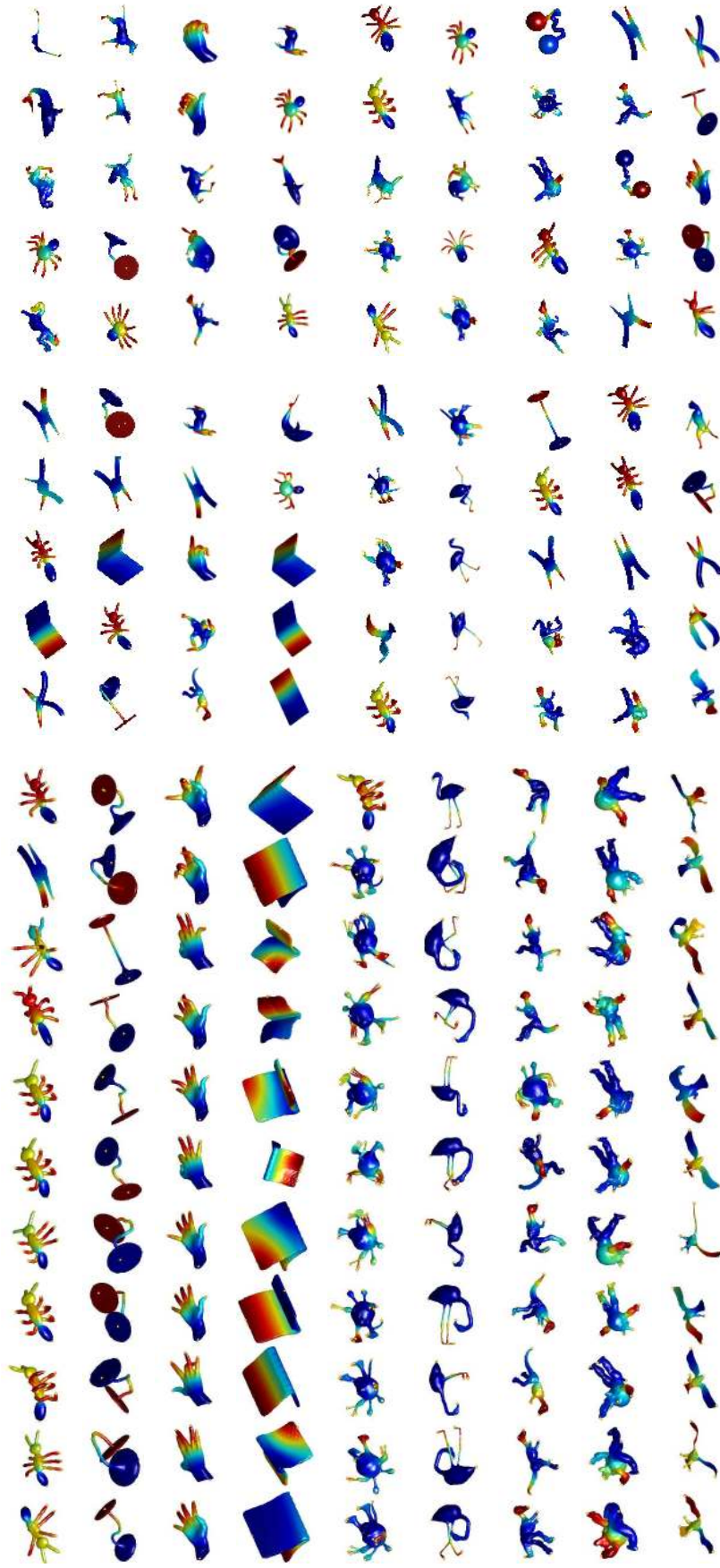


Figure 4.6: Shape retrieval results of SHREC 2011 data set. Left: queries. Middle: First 15 matches using the SRCP-TD descriptor. Right: matches from 30 to 35. The color represents the mean temperature induced from only the detected critical points shown in yellow.

Method	NN	1-Tier	2-Tier	e-Measure	DCG
TD	0.59	0.36	0.50	0.33	0.66
Shape-DNA	0.96	0.64	0.87	0.55	0.90
SRCP-TD	0.96	0.67	0.90	0.56	0.91
TD	0.25	0.18	0.31	0.20	0.503
Shape-DNA	0.94	0.78	0.90	0.55	0.93
SRCP-TD	<b>0.97</b>	<b>0.81</b>	<b>0.93</b>	<b>0.56</b>	<b>0.95</b>

Table 4.1: Results on the non-rigid world data set. Upper four rows are results using KNN, while lower four rows are results using CRC-RLS.

Method	NN	1-Tier	2-Tier	e-Measure	DCG
TD	0.64	0.37	0.47	0.33	0.66
Shape-DNA	0.99	0.85	0.92	0.67	0.96
SRCP-TD	<b>0.99</b>	<b>0.86</b>	<b>0.94</b>	<b>0.69</b>	<b>0.96</b>
TD	0.18	0.14	0.24	0.16	0.46
Shape-DNA	0.56	0.58	0.74	0.55	0.78
SRCP-TD	<b>0.97</b>	<b>0.80</b>	<b>0.89</b>	<b>0.65</b>	<b>0.94</b>

Table 4.2: Results on SHREC’11 data set. Upper four rows are results using KNN, while lower four rows are results using CRC-RLS.

This experiment also investigated the effect of the reduction of the SRCP-TD descriptor on the performance on the SHREC 2011 data set. Table 4.4 lists the five evaluation measures for several sizes of the descriptor. Expectedly, the performance improves as the size of the descriptor is increased. However after a size of 104 (or even 64), the improvement is rather marginal. So a compact descriptor of size 64 can be sufficient to provide high performance on the data set.

Method	NN	1-Tier	2-Tier	e-Measure	DCG
SRCP-TD	<b>0.9833</b>	0.7621	0.9209	0.5633	<b>0.9493</b>
SRCP-TD 64	0.9750	<b>0.8056</b>	<b>0.9278</b>	<b>0.5652</b>	0.9480

Table 4.3: Results on the non-rigid world data set using the proposed descriptors computed at only one critical point. Upper row using KNN, while lower row using CRC-RLS.

N	NN	1-Tier	2-Tier	e-Measure	DCG
24	0.7733	0.6838	0.8175	0.5886	0.8653
44	0.9200	0.7362	0.8509	0.6173	0.9059
<b>64</b>	<b>0.9600</b>	<b>0.7783</b>	<b>0.8818</b>	<b>0.6446</b>	<b>0.9323</b>
84	0.9683	0.7944	0.8875	0.6496	0.9385
<b>104</b>	<b>0.9783</b>	<b>0.8087</b>	<b>0.8976</b>	<b>0.6592</b>	<b>0.9453</b>
124	0.9783	0.8111	0.8998	0.6602	0.9467
144	0.9783	0.8106	0.8996	0.6603	0.9468

Table 4.4: Results on SHREC'11 data set with respect to size (N) of the SRCP-TD descriptor using the CRC-RLS.

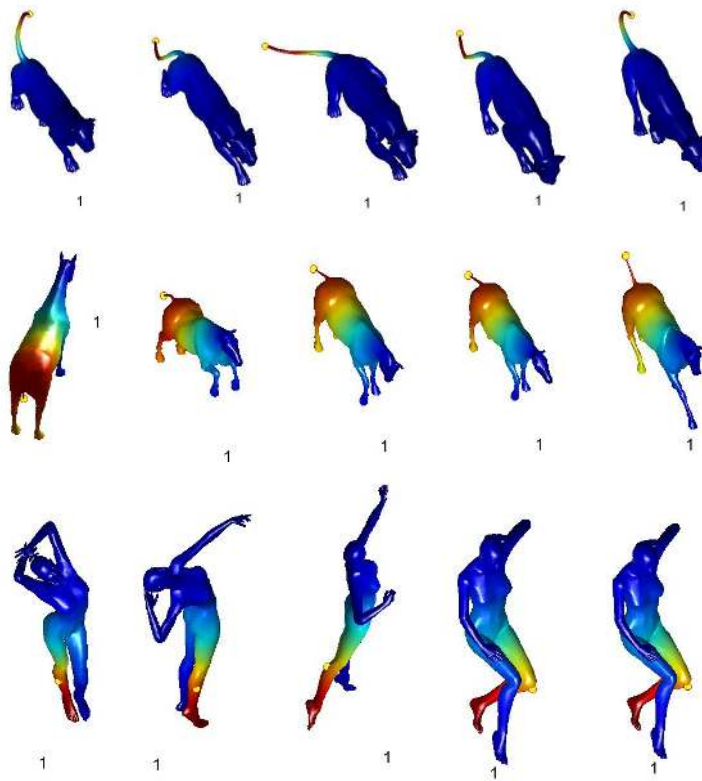


Figure 4.7: Single critical points detected for several shapes of the non-rigid world data set.

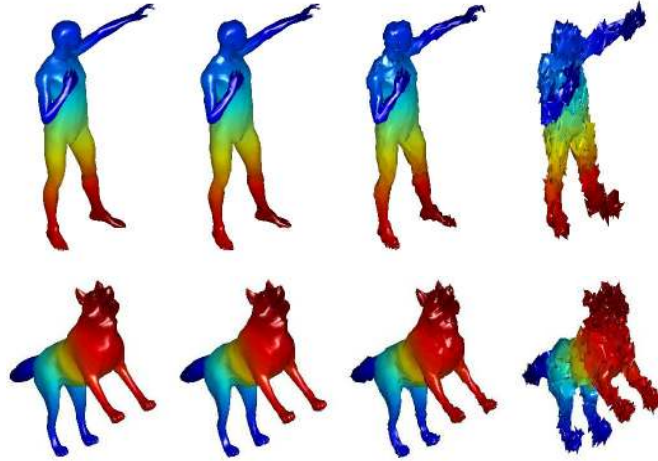


Figure 4.8: noise-corrupted models. from left to right 0.0, levels 1, 2, and 3  $NSR$ .

#### 4.2.4 Noise Effect Analysis

Normally 3D smoothing filter is used to preprocess the data before calculating the HK. Thus this method is expected to be less sensitive to noise. To demonstrate this, an experiment is done on non-rigid world data set by perturbing the vertices of the original model using various levels of numerical noise. By adding Gaussian noise with varying Noise-to-Signal Ratio ( $NSR$ ) between the variance of noise and variance of the original signal (coordinates of the vertices), the number of critical points for the centaur shape is 22 at  $NSR = 0.001$ , and 22 at  $NSR = 0.01$  versus 19 for the noise-free shape, which shows rather good insensitivity to noise. Figure 4.8 shows sample 3D models, corrupted by different levels of noise (before the filter step). Table 4.5 gives the five evaluation measures on the non-rigid world data set computed for the proposed descriptor computed at different  $NSR$ . The tabulated measures also demonstrate that the proposed descriptor is rather insensitive to noise.

NSR	NN	1-Tier	2-Tier	e-Measure	DCG
0.0	0.97	0.81	0.93	0.56	0.95
0.001	0.92	0.74	0.92	0.55	0.93
0.01	0.87	0.67	0.85	0.50	0.88

Table 4.5: Results on non-rigid world data set with different NSR of the SRCP-TD descriptor using the CRC-RLS.

### 4.3 A Shape Descriptor Based on Scale Invariant Heat Kernels

This section develops a new feature descriptor based on scale-invariant heat kernels [31]. The proposed method uses the first non-trivial Laplace-Beltrami eigenfunction to detect a small number of sparse critical points on the shape surface, as described in Section 4.1. These points are robust to the shape class, and their number can in itself be used as one of the discriminatory features among the various classes. Then the HK is calculated only at the detected critical points at different time scales. This reduces the size or dimensionality of the descriptors, yet does not sacrifice the accuracy of the presentation. Then scale invariance is achieved using a novel transformation method that is simpler and considerably less noise sensitive than the method proposed in [3]. A concatenation of the histograms of the significant components of the scale-invariant HK for all the points is used as a feature vector for classification. The resulting descriptor is then used for classification via the relaxed collaborative representation (RCR) model [96, 97]. The resulting descriptor captures the local as well as global shape information since it uses the temperature distribution at the critical points at several time samples.

This section also evaluates the proposed descriptor on the problem of 3D shape retrieval and object classification. For the sake of comparison, the approach is compared to the state-of-the-art approaches, namely, the shape-DNA [2], the



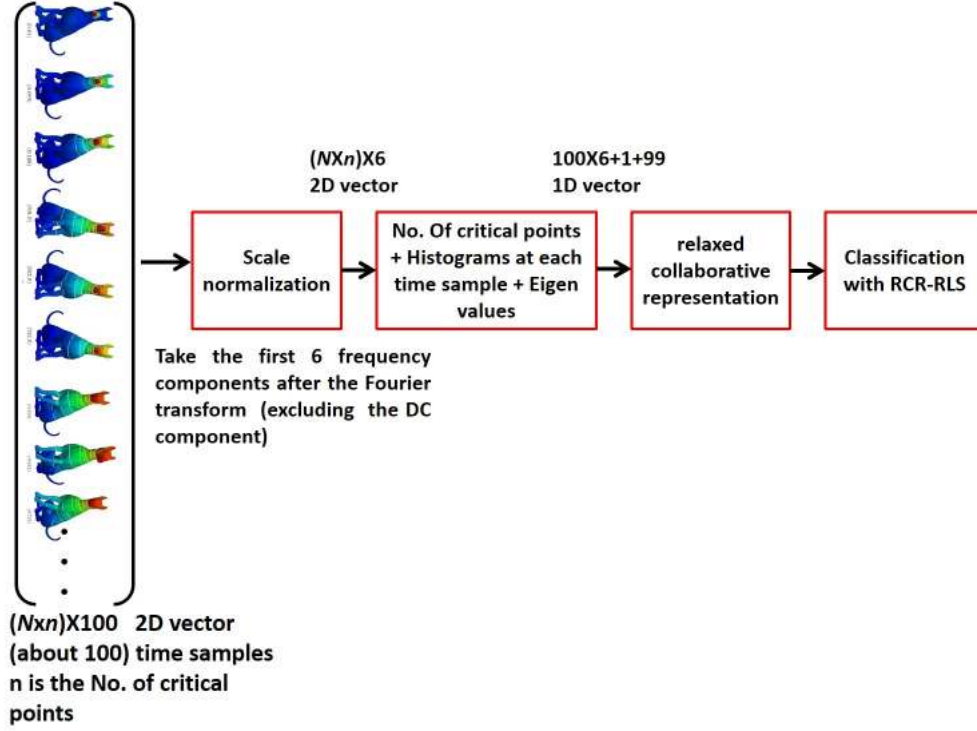


Figure 4.9: Construction of the proposed descriptor ( $N$  is the number of object vertices).

TD approach [1] and the Shape Google approach [3] on two standard 3D data sets, the nonrigid world data set [5] and the SHREC 2011- Shape Retrieval Contest of Non-rigid 3D Watertight Meshes [98]. Demonstration shows that the new approach can outperform these methods even on different object scales and considerably noisy data.

### 4.3.1 Weak And Discrete Formulation of The Heat Equation

To compute the solution to (3.13) the weak formulation through the Galerkin formulation is used, as described in Section 3.4.2. Then the heat kernel will be:

$$K(\mathbf{x}, \mathbf{y}, t) := \phi D(t) \phi^T B, \quad (4.2)$$

### 4.3.2 Scale Invariance

Scale invariance is a desirable property of the shape descriptor, which can be achieved by one of the following four different methods:

1. Trying detecting the scale, as done in most feature descriptors (e.g. SIFT) in image analysis. However, 3D shapes are usually poorer in features and scale detection can be done reliably only at a sparse set of feature points.
2. Normalizing of Laplace-Beltrami eigenvalues, but this method may suffer if the object has missing parts [75]. In such case, the scale invariance must be introduced locally rather than globally.
3. Utilizing a series of transformations applied to the HKS [75] in order to avoid scale detection. This allows creating a dense descriptor. This method is considered local, thus can work with objects with missing parts.
4. Using the local equi-affine invariant Laplace-Beltrami operator proposed by [79].

This work, proposes a novel local scale normalization method based on simple operations (thus belonging to the third category above). It was shown [75] that scaling a shape by a factor  $\beta$  results in changing  $H(\mathbf{x}, t)$  to  $\beta^2 H(\mathbf{x}, \beta^2 t)$ .

Thus, a series of transformations are applied to HK as follows. Starting from each critical point  $\mathbf{x}$ , the HK is sampled at every surface point  $\mathbf{y}$  algorithmically in time ( $t = \alpha^\tau$ ) and the function:

$$h_\tau = H(\mathbf{x}, \alpha^\tau) \quad (4.3)$$

is formed. Scaling the shape by  $\beta$  results in a time shift  $s = 2 \log_\alpha \beta$  and amplitude scaling by  $\beta^2$ . That is:

$$h'_\tau = \beta^2 h_{\tau+s} \quad (4.4)$$

[75] proposed to take the logarithmic transformation  $\log h'_\tau = 2 \log \beta + \log h_{\tau+s}$  which decouples the multiplicative constant from  $h_{\tau+s}$ . Then they proposed to take the derivative afterwards to remove the effect of the resulting additive  $2 \log \beta$  term and then taking the amplitude of the Fourier Transform (FT) of the derivative to remove the effect of the time shift  $s$ . Since the derivative operator is sensitive to noise, their method is not robust enough. This work proposes to apply the FT directly to  $h'_\tau$  in Equation 4.4 to get:

$$H'(w) = \beta^2 H(w) \exp(j 2\pi w s). \quad (4.5)$$

Then taking the amplitude of the FT it will be:

$$| H'(w) | = \beta^2 | H(w) | \quad (4.6)$$

The effect of the multiplicative constant  $\beta^2$  is eliminated by normalizing the  $| H'(w) |$  by the sum of the amplitudes of the FT components. The amplitudes of the first significant FT components (normally 6) are employed to construct the scale-invariant shape descriptor. This proposed method eliminates the scale effect without having to use the noise-sensitive derivative operation or the logarithmic transformation that both were used in [75]. Thus, the proposed method is simpler, more computationally-efficient and more robust to noise. This is clearly verified in Figure 4.10 that shows the scale-invariant heat kernel for a HK computed at a vertex of a 3D shape and another HK computed for the same object with 3-times scale up under different noise levels. The two descriptors computed at the two different scales are virtually identical using the proposed method even at high noise levels. The method in [3] demonstrates a significant influence of the noise on the computed descriptors.

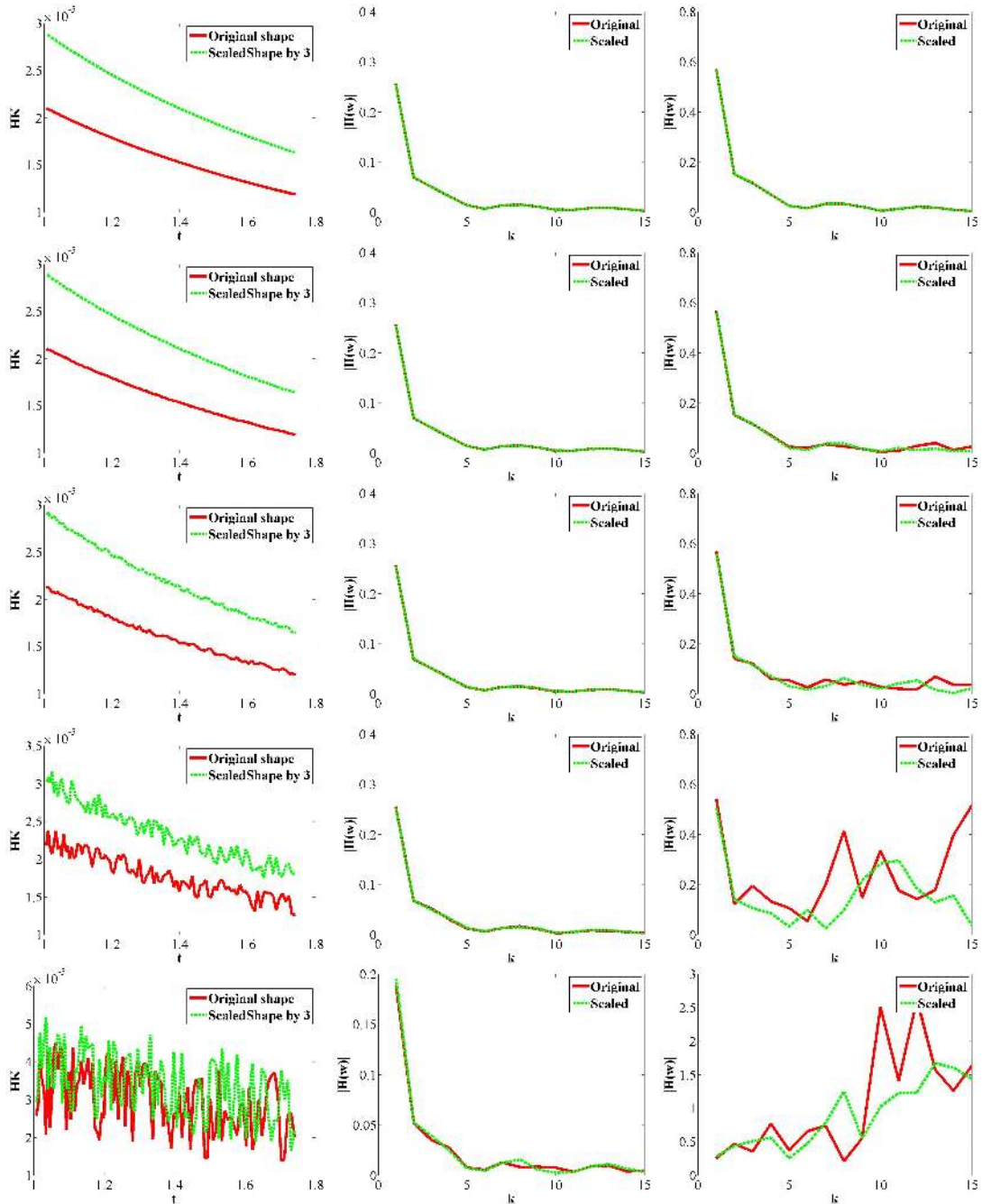


Figure 4.10: Construction of the scale-invariant heat kernel under several noise levels. (Left) heat kernel computed at a point on a shape at different time (red) and the scaled heat kernel (in green) computed at a corresponding point on a shape scaled up by a factor of 3. (Middle) The amplitude of the first 15 fourier transform components of  $|H(w)|$  for the two heat kernels (again in red and green) using the proposed method showing complete overlapping specially in the first four noise levels. (Right) Similar but using the method in [3]. First row shows signal without noise. Then noise level is increased in the subsequent rows.

### 4.3.3 Scale Invariant Heat Kernel Descriptor

Figure 4.9 shows the steps for the construction of the proposed descriptor. The proposed shape descriptor is constructed as follows: HKs are calculated at some critical points detected on the surface at various (about 100) time samples. The dimension of the heat kernel is  $N \times 100 \times n$  where  $N$  is the number of the shape vertices, and  $n$  is the number of detected critical points. Then scale-invariance is introduced in the computed HK as explained in the previous subsection. The first 6 FT component are used at each critical point. The size of the feature now is  $N \times 6 \times n$ . Since it would be difficult to compare descriptors of shapes with different number of vertices, histograms of scale-invariant HK is used to overcome the descriptor alignment problem and to reduce the descriptor size. For each FT component, a histogram of 100 pins for all HKs is calculated. This reduces the descriptor size to  $100 \times 6$  1D feature vector. Afterwards, the number of the detected critical points is added to this descriptor. In addition, the normalized eigenvalues of the Laplace-Beltrami operator are appended. Eventually, the size of the feature vector is 700. This vector, dubbed Critical Points-based Heat Kernel (CP-HK), can be used for classification using some well-known classifiers. However, for the latter part, the collaborative classification [96] is used as a feature classifier.

### 4.3.4 Collaborative Classification

Recently, collaborative representation has also been used in pattern classification. Zhang [97] proposed a new classification scheme, namely collaborative representation (CR) based classification with regularized least square (CRC-RLS), which has significantly less complexity than SRC but leads to very competitive classification results. Then [96] propose a relaxed collaborative representation (RCR) model. Zhang [96] showed that RCR is simple, and very competitive with

state-of-the-art image classification methods.

It considers both the similarity and distinctiveness of different features in coding and classification stages.

A new term is added to the collaborative representation equation to regularize the coding vectors of different features over their associated dictionaries

$$\min_{\alpha_k} \sum_{k=1}^K \omega_k \|\alpha_k - \bar{\alpha}_k\|_2^2 \quad (4.7)$$

where  $\alpha_k, k = 1, 2, \dots, K$  is the coding vector of the  $k^{th}$  feature vector  $y_k$  over the  $k^{th}$  dictionary  $D_k$ ,  $\bar{\alpha}_k$  is the mean vector of all  $\alpha_k$ , and  $\omega_k$  is the weight assigned to the  $k^{th}$  feature. So the collaborative representation equation will be

$$\min_{\alpha_k} \sum_{k=1}^K (\|y_k - D_k \alpha_k\|_2^2 + \lambda \|\alpha_k\|_2^2 + \tau \omega_k \|\alpha_k - \bar{\alpha}_k\|_2^2) \text{ s.t. } \text{prior} \omega_k \quad (4.8)$$

where  $\lambda$  and  $\tau$  are positive constants and *prior* $\omega_k$  means the prior made on weights  $\omega_k$ . The weights are optimized simultaneously with the coding to address the distinctiveness of different features. In the classification stage, the query sample is assigned to the class which yields the lowest weighted coding residual. The RCR approach is used for the coding and classification of the proposed descriptors. More details about the algorithm in [96].

### 4.3.5 Experimental Results

To test the performance of the proposed descriptors two data sets are used. The first one is the nonrigid world data set [4, 5]. The database contains a total of 148 3D nonrigid objects in a variety of poses, including 9 cats, 11 dogs, 3 wolves, 17 horses, 15 lions, 21 gorillas, 1 shark, 24 female figures, and two different male figures, containing 15 and 20 poses. The database also contains 6 centaurs, and 6 seahorses for partial similarity experiments. Each object contains approximately

3400 vertices. The second database is the SHREC 2011 - Shape Retrieval data set [98]. This is a large-scale database consists of 600 non-rigid 3D objects that are derived from 30 original models.

For the sake of comparison, this work shows the results of the Shape-DNA approach [2], describing shapes by the vector of the first eigenvalues of the Laplace-Beltrami operator. The first 15 eigenvalues are used to construct the Shape-DNA descriptors. Eigenvalues were computed using the same cotangent weight discretization. Also the proposed approach is compared to the method in [1] that uses the (TD) as a shape descriptor with an  $L_2$  classifier, and the Shape Google approach [3] that is based on the bag of features (BoF) technique.

Figure 4.11 shows sample shape retrieval results of the CP-HK descriptor on the nonrigid world data set using the RCR classifier. The retrieved objects per each query are ranked from left to right based on the distance measure of the classification algorithm. The figure clearly demonstrates the high retrieval rate of the approach. It is interesting to note that the objects similar to the query pose are retrieved first, then come objects of the same shape but with different poses. This shows that this descriptor can be used for object pose estimation as well. The wrong matches are shown in light green, where some of them are only retrieved after all objects of the correct shape have been already retrieved.

Figure 4.12 shows sample shape retrieval results of the CP-HK descriptor on the SHREC 2011 data set. The figure shows the first 20 matches for each query ranked according to the distance measure of the RCR classifier. Note that these first ranked matches cover almost all the shapes found in the database belonging to the same query shape, which indicates the high quality of the proposed shape descriptor.

An illustration of confusion matrix for the SHREC'11 data set using the CP-HK descriptor is shown in Figure 4.13. It is clear that the confusion matrix is diagonal, and almost the first 20 matched shapes for each query come are from

Measure	TD	Shape-DNA	Shape-Google	Proposed method
NN	0.5917	0.9833	0.9000	<b>0.9667</b>
1-Tier	0.3646	0.7972	0.6666	<b>0.7750</b>
2-Tier	0.5036	0.9298	0.8656	<b>0.9053</b>
e-Measure	0.3307	0.5618	0.5443	<b>0.5430</b>
DCG	0.6654	0.9585	0.8874	<b>0.9389</b>

Table 4.6: Results on non-rigid world data set. The proposed method is compared to TD [1], Shape-DNA [2], and Shape-Google [3].

Measure	TD	Shape-DNA	Shape-Google	Proposed method
NN	0.6483	0.9833	0.9567	<b>0.9733</b>
1-Tier	0.3704	0.6917	0.6225	<b>0.7798</b>
2-Tier	0.4768	0.8050	0.7288	<b>0.8823</b>
e-Measure	0.3369	0.5827	0.5245	<b>0.6443</b>
DCG	0.6684	0.9155	0.8718	<b>0.9364</b>

Table 4.7: Results on SHREC’11 data set. The proposed method is compared to TD [1], Shape-DNA [2], and Shape-Google [3].

the same class.

For the sake of quantitative assessment of the approach performance, the following standard five evaluation measures are recorded (see [100] for detailed definitions): Nearest Neighbor (NN) where  $N = 1$ , First Tier (FT), Second Tier (ST), E-measure (E), and Discounted Cumulative Gain (DCG). Table 4.6 lists these measures on the nonrigid world data set, and Table 4.7 lists these measures on the the SHREC 2011 data set. The tables compares the proposed approach against the TD, Shape-DNA and Shape Google approaches. The proposed approach significantly outperforms the TD and the Shape Google approaches on both data sets. While the Shape-DNA approach shows slightly better performance than the proposed approach on the non-rigid world data set, the performance measures become in favor of the proposed approach on the larger-scale SHREC’11 database. Moreover, the Shape-DNA performance severely suffers on noisy and scaled data, as will be demonstrated in the next experiment.





Figure 4.11: Some shape retrieval results on the nonrigid world data set. Left: Queries. Right: First 10 matches using the CP-HK descriptor. The detected sparse points are shown in yellow. The light green-colored shapes are wrong matches.

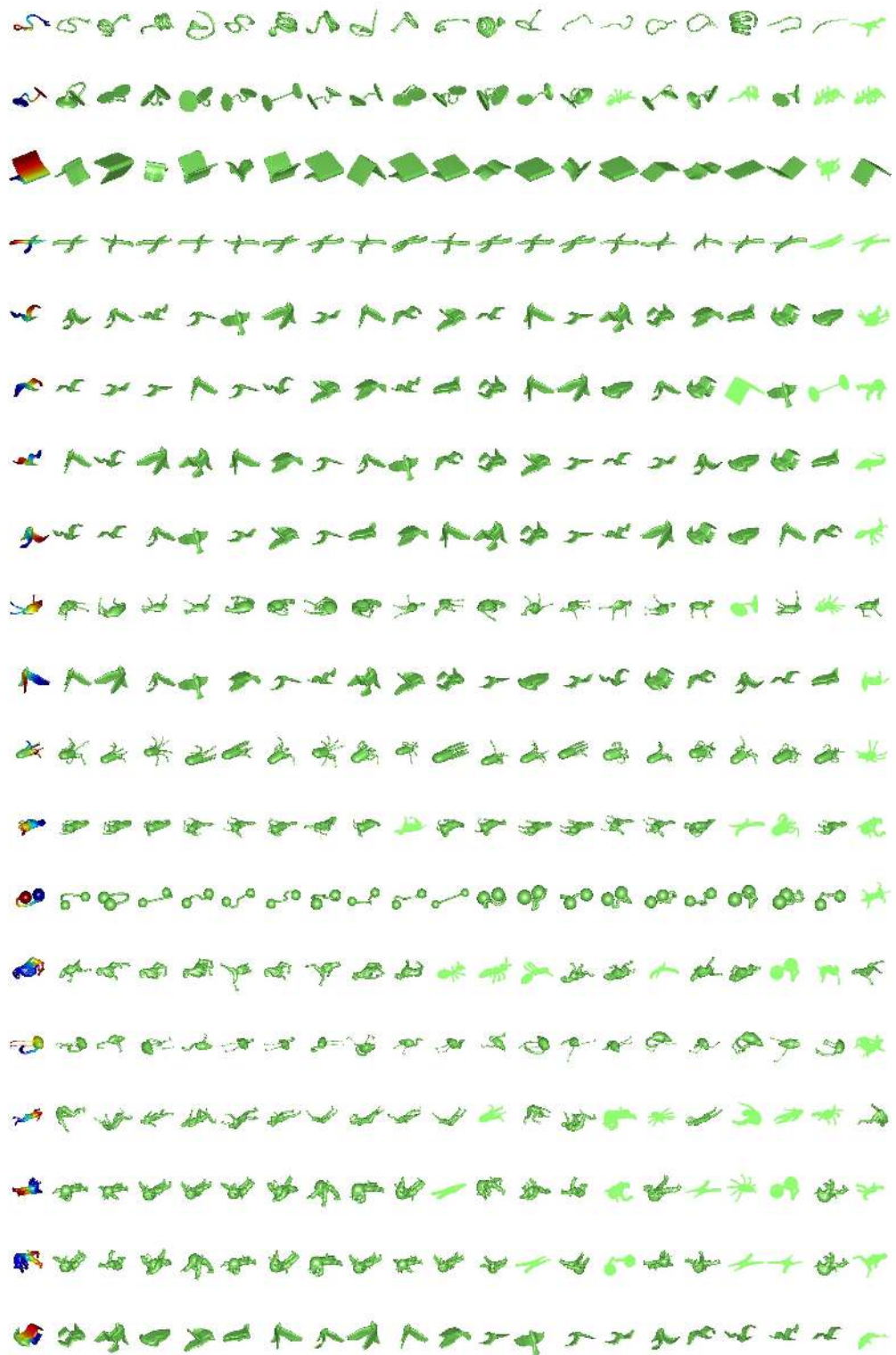


Figure 4.12: Shape retrieval results of SHREC'11 data set. Left: queries (colored). Then First 20 matches using the proposed descriptor. The light green colored shapes represent wrong matches. The detected critical points are shown in yellow.

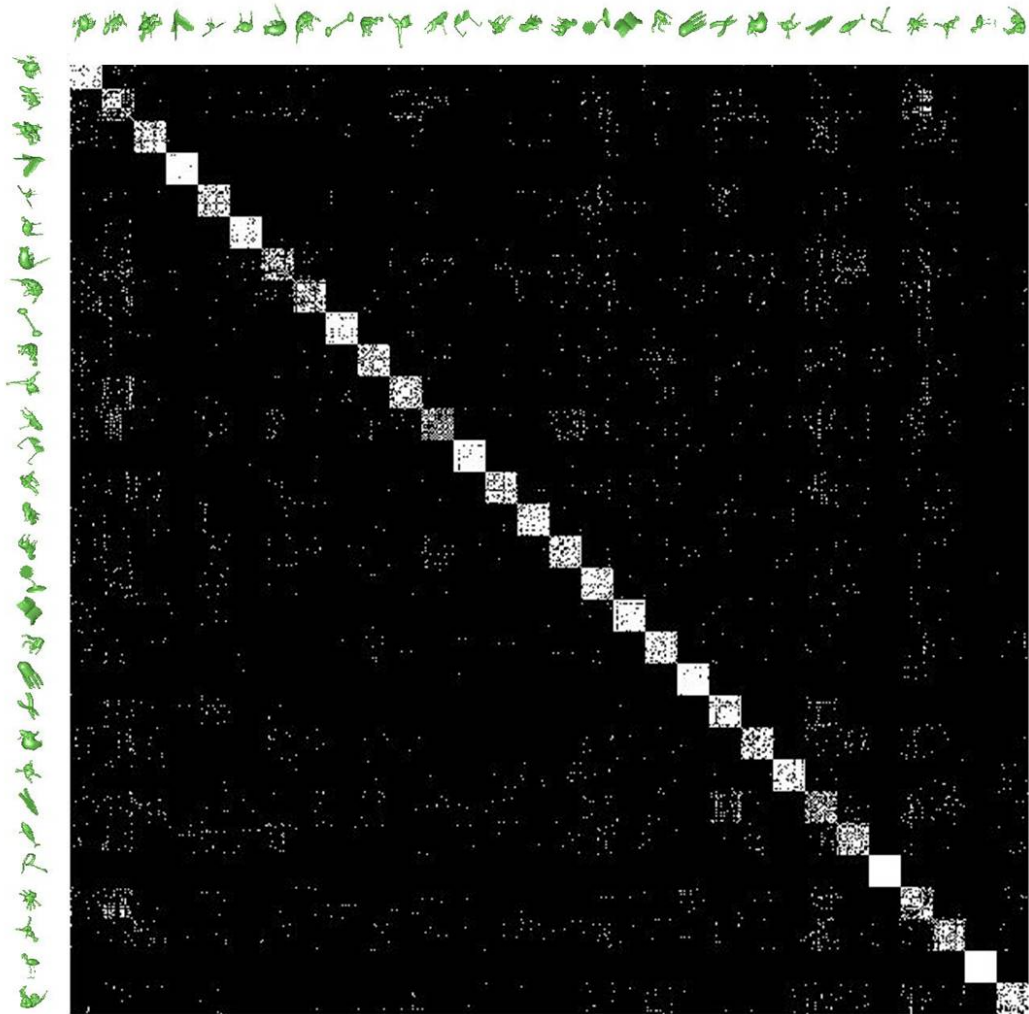


Figure 4.13: An illustration of confusion matrix ( $600 \times 600$ ) for the SHREC'11 data set. Each row in this matrix represents the first 20 match retrieval results for each shape in the data set(600 shapes). Note that the shapes belonging to the same class are sorted to be next to each other.

White Noise Level	Proposed Approach	Shape Google	Shape-DNA
1	1.0000	1.0000	1.0000
2	1.0000	0.9000	0.9667
3	0.9000	0.1333	0.6000
Shot Noise Level	Proposed Approach	Shape Google	Shape-DNA
1	1.000	0.93331	1.000
2	1.000	0.86662	0.9667
3	0.966	0.53333	0.4000
Scale Level	Proposed Approach	Shape Google	Shape-DNA
1	1.0000	0.8000	0.9667
2	1.0000	0.4666	0.9667
3	1.0000	0.2333	0.9667

Table 4.8: Performance versus white noise, shot-noise, and scale in three severity levels of the proposed CP-HK descriptor using the RCR classifier compared to those of Shape Google [3] and Shape-DNA [2] (1.00 mean 100%).

### 4.3.6 Distorted Data Experiment

Another experiment is carried out to assess the approach performance under several distorted data scenarios. Here the performance of the proposed approach is compared with those of the Shape-DNA and Shape Google approaches. A query set is formed consisting of 30 shapes taken from the SHREC'11 data set, after applying several distortions: a Gaussian white noise, shot-noise, and different scales. The performance in terms of the Nearest Neighbor (NN) measure [100] versus these three distortions in three different levels of severity is shown in Table 4.8. Figure 4.14 illustrates sample shapes corrupted with three degrees of white noise along with the query results using the proposed approach. Also Figure 4.15 illustrates sample shapes corrupted with three degrees of shot noise along with the query results using the proposed approach. The performance against missing parts is demonstrated in Figure 4.16.

The proposed approach has shown significantly better results than the other two approaches. For example, the proposed approach has 90% accuracy on the third severe level of white noise, whereas the Shape Google performance drops

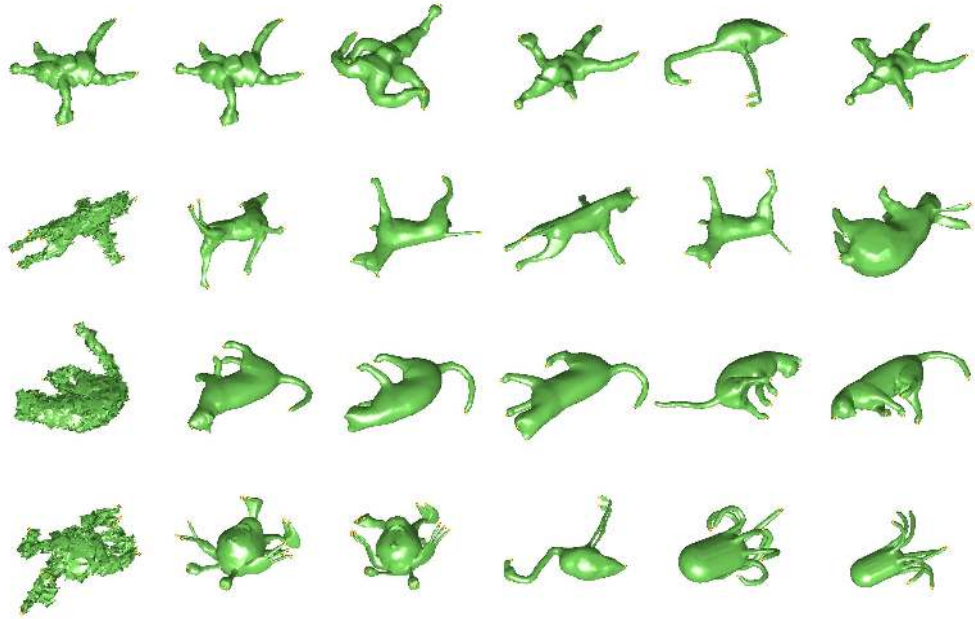


Figure 4.14: Some shape retrieval results for different shapes with different noise levels. Left: queries. Right: the first 5 matched shapes.

to 13% and Shape-DNA has only 60% accuracy. Moreover, the proposed approach has retrieved shapes with different scales with 100% accuracy, thanks to the new scale invariance method, while the Shape Google presented considerably lower performance. The Shape-DNA has a constant performance against scale due to its normalization of the Laplace-Beltrami eigenvalues. Note that in this experiment, in contrast to a similar Shape Google experiment reported in [3], similar-class positive shapes (males and females, centaur, horse, and human shapes) did not ignored. As such, a result of the female shape in response to a male shape query, for example, is counted against the method, which was not in [3]. This justifies why the performance of Shape Google in Table 4.7 and Table 4.8 is lower compared to that reported in [3].

### 4.3.7 Using Different Classifiers

Without combining the *normalized* eigenvalues of the Laplace-Beltrami operator with the scale invariant HK descriptor, several classifiers are experi-

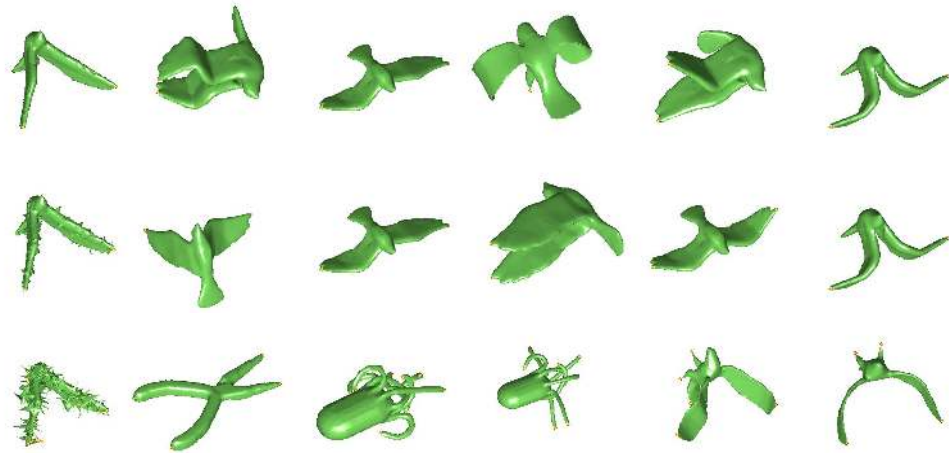


Figure 4.15: Some shape retrieval results for different shapes with different shot noise levels. Left: queries. Right: the first 5 matched shapes.

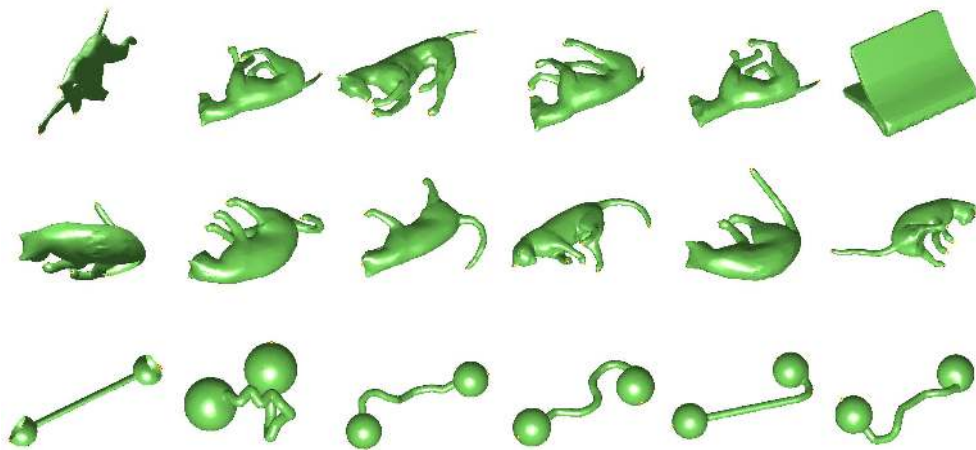


Figure 4.16: Some shape retrieval results for different shapes with different missing parts. Left: queries. Right: the first 5 matched shapes.

Classifier	TP rate	FP rate	Precision	Recall	ROC area
C4.5	0.883	0.016	0.893	0.883	0.939
Naive Bayes	0.858	0.022	0.845	0.858	0.954
NN	0.883	0.021	0.901	0.883	0.931
SVM	0.908	0.015	0.917	0.908	0.984
C4.5	0.608	0.067	0.601	0.608	0.789
Naive Bayes	0.542	0.071	0.57	0.542	0.848
NN	0.533	0.079	0.535	0.533	0.727
SVM	0.642	0.063	0.641	0.642	0.876
C4.5	0.933	0.009	0.94	0.933	0.962
Naive Bayes	0.867	0.018	0.876	0.867	0.983
NN	0.975	0.005	0.975	0.975	0.985
SVM	0.983	0.003	0.983	0.983	0.996

Table 4.9: Average accuracy measures on the non-rigid world data set with different classifiers. For all measures, the higher the better except the FP rate, where the lower the better. Upper four rows are the proposed approach, middle are for the TD method, while the lower four rows are for the Shape-DNA.

mented [101], namely, the Nearest Neighbor (NN) classifier, Naive Bayes classifier, Decision trees using the C 4.5 algorithm, and Support Vector Machines (SVM).

For the sake of quantitative assessment of the approach performance with all the tested classifiers, the following standard five evaluation measures are recorded: True Positive (TP) rate, False Positive (FP) rate, precision, recall, and ROC area. Table 4.9 lists the average measures over all classes in the non-rigid world data set. The table compares the proposed descriptor against the TD and the shape-DNA descriptors using the four different classifiers. Analogously, Table 4.10 shows the performance on the SHREC 2011 data set. Except for the Shape-DNA, the CP-HK descriptor exhibits the best performance using the SVM. Although the Shape-DNA shows very good performance in this experiment, (except for the SHREC data set with the SVM classifier it is not good) it severely suffers when there are missing parts in the objects (i.e., on partial shape matching). This is clearly demonstrated in the next experiment.

Another experiment is carried out to assess the approach performance under several distorted data scenarios. A query set consisting of 30 shapes are

Classifier	TP rate	FP rate	Precision	Recall	ROC Area
C4.5	0.826	0.006	0.827	0.826	0.916
Naive Bayes	0.859	0.005	0.885	0.859	0.974
NN	0.862	0.005	0.88	0.862	0.929
SVM	0.914	0.003	0.916	0.914	0.983
C4.5	0.646	0.012	0.659	0.646	0.831
Naive Bayes	0.648	0.012	0.69	0.648	0.922
NN	0.513	0.017	0.569	0.513	0.748
SVM	0.619	0.013	0.596	0.619	0.925
C4.5	0.951	0.002	0.954	0.951	0.975
Naive Bayes	0.966	0.001	0.967	0.966	0.998
NN	0.99	0	0.99	0.99	0.995
SVM	0.824	0.006	0.873	0.824	0.978

Table 4.10: Average accuracy measures on the SHREC'11 data set with different classifiers. For all measures, the higher the better except the FP rate, where the lower the better. Upper four rows are the proposed approach, middle are for the TD method, while the lower four rows are for the Shape-DNA.

Noise Level	TP rate	FP rate	Precision	Recall	ROC area
1	0.933	0.002	0.900	0.933	0.998
2	0.900	0.003	0.85	0.900	0.996
3	0.733	0.009	0.636	0.733	0.948
1	0.800	0.007	0.733	0.800	0.756
2	0.800	0.007	0.733	0.800	0.756
3	0.733	0.009	0.644	0.733	0.672

Table 4.11: Average accuracy measures on the SHREC'11 data set with different noise levels, Upper is the proposed method, Lower is Shape-DNA.

formed from the SHREC'11 data set, after applying several distortions: a Gaussian white noise, shot-noise, scaling each in three different levels of severity. The performance versus noise and shot noise is shown in Table 4.11 and Table 4.12. Shapes with different scales have been retrieved with 100% accuracy at rank 1 as shown in Table 4.13.

## 4.4 Summary

This chapter presented two approaches for shape matching and retrieval based on heat kernel (HK), and scale-invariant heat kernel (HK). For the first



Noise Level	TP rate	FP rate	Precision	Recall	ROC area
1	0.800	0.007	0.722	0.800	0.9447
2	0.867	0.005	0.825	0.867	0.958
3	0.467	0.018	0.394	0.467	0.928

Table 4.12: Average accuracy measures on the SHREC’11 data set with different shot-noise levels.

Scale Level	TP rate	FP rate	Precision	Recall	ROC area
1,2,3	1	0	1	1	1
1,2,3	0.833	0.006	0.767	0.833	0.971

Table 4.13: Average accuracy measures on the SHREC’11 data set with different scale levels. Upper is the proposed method, Lower is Shape-DNA.

approach proposed to use the first non-trivial Laplace-Beltrami eigenfunction to detect a small number of sparse critical points on the surface of the shape. These points were shown to be robust to the shape class, and their count can in itself be used as one of the discriminatory features among the various classes.

A feature vector (descriptor) computed at these critical points is sufficient for shape matching. Sparse representation of the descriptor can be obtained to reduce its dimensionality before shape classification. The proposed descriptors are compact in size, the feature vector can be down to 64, and efficient in computation since it is constructed at sparse points.

The experimental results have shown that the proposed descriptor can achieve state-of-the-art results on the two benchmark data sets. The results have indeed confirmed the improved performance of the sparse representation over the original non-reduced descriptors, yet reducing the time and space complicity for the shape retrieval problem. In addition, results have shown that the descriptor can be constructed using one critical point with no significant sacrifice in the performance.

The second approach presented a new approach for shape matching and retrieval based on scale-invariant heat kernel (HK). The first non-trivial Laplace-

Beltrami eigenfunction is used to detect a small number of sparse critical points on the surface of the shape. These points were shown to be robust to the shape class, and their count can in itself be used as one of the discriminatory features among the various classes. The proposed descriptor is based on the histograms of HKs at different time scales computed at these critical points. It thus can capture the local (at small time scales) as well as global shape information (at bigger time scales). A new local transformation method to introduce scale-invariance has been also proposed, which is shown to be simpler, more computationally efficient and noise-robust compared to a well-known method [3]. A collaborative classification scheme for object matching and retrieval is used to classify the proposed descriptor.

The experimental results have shown that the proposed descriptor can achieve high performance on two popular benchmark data sets. The proposed approach has outperformed state-of-the-art approaches for shape representation and retrieval. The results have demonstrated that the proposed descriptor enjoys many of the properties desirable in a shape descriptor (those given in Chapter 1.4): It is isometry-invariant, able to deal with the non-rigid shape deformation. It is independent of the object's size. It is stable to small changes in the shape. It is less sensitive to noise, as the descriptor is computed at a number of points on the shape, rather than the whole shape. It is computationally efficient. And it has similarity, uniqueness and scope properties, it has retrieved almost all the database shapes truly matching the same query shape. An important observation from experiment results is that the proposed approach is more able to handle data under several distortion scenarios (white noise, shot-noise, and scale changes) than the well-known Shape Google and Shape-DNA approaches.

## CHAPTER 5

# HEAT DIFFUSION OVER WEIGHTED MANIFOLDS: A NEW DESCRIPTOR FOR TEXTURED 3D NON-RIGID SHAPES

This chapter proposes an approach for modeling textured 3D non-rigid models based on Weighted Heat Kernel Signature(W-HKS). As a first contribution, this chapter shows how to include photometric information as a weight over the shape manifold, it also proposes a novel formulation for heat diffusion over weighted manifolds. As a second contribution this chapter presents a new discretization method for the proposed heat equation using finite element approximation. Finally, the weighted heat kernel signature is used as a shape descriptor. The proposed descriptor encodes both the photometric, and geometric information based on the solution of one equation. The performance is tested on a benchmark data set. The results have indeed confirmed the high performance of the proposed approach on the textured shape retrieval problem, and show that the proposed method is useful in coping with different challenges of shape analysis where pure geometric and pure photometric methods fail.

### 5.1 Introduction

Recently, many sensors are able to acquire the color information besides the 3D shape, also multiple-view stereo techniques are able to recover both geometric and photometric information. These photometric features can play an

important role in many shape analysis applications, such as shape matching and correspondence because it contains rich information about the visual appearance of real objects. This new requirement and its important applications adds another, new dimension to the problem difficulty. Most descriptors proposed so far are confined to shape, that is, they analyze only geometric and/or topological properties of 3D models. Therefore more efforts are needed to be done to consider color in addition to shape in object representation and description. The domain of research in this chapter is the representation of textured shapes in order to develop an efficient descriptor that combines the color information as well as the geometric shape information. The sought representation should cope with non-rigid transformations, which is a key requirement for many target applications.

### 5.1.1 Review of Related Work

There has been an extensive work on constructing descriptors for 3D shapes (e.g., [15, 16, 19, 22, 23, 68, 69]). One of the challenging issues in that regard is how to handle handling non-rigid transformation. The problem of non-rigid shape deformation needs more work to compensate for the degrees of freedom resulting from local deformations. In the past decade, significant effort has been invested in extending the invariance properties to non-rigid shape deformations.

Sun et al. [74], and [102] proposed heat kernel signatures (HKS) as deformation-invariant descriptors based on diffusion of multi-scale heat kernels. HKS is a point based signature satisfying many of the good descriptor properties but suffers from sensitivity to scale. The authors did not demonstrate how to retrieve shapes using HKS, although they pointed out the future potentials in shape retrieval applications. Bronstein et al. [3, 75] and the current dissertation [31] solved the HKS scale problem through a series of transformations.

All these efforts have focused only on the 3D shape. Recently, taking the photometric information into account to calculate a 3D shape descriptor has at-

tracted more research. Liu et al. [103] proposed a method that picks points in regions of either geometry-high variation or color-high variation, and defines a signature at these points. A geometric SIFT-like descriptor for textured shapes defined directly on the surface in [104]. The work of Kovnatsky et al. [29] is the most related work to the proposed approach; it uses the diffusion geometry framework for the fusion of geometric and photometric information in local and global shape descriptors. Their construction is based on an *ad hoc* definition of a diffusion process on the shape manifold embedded into a high-dimensional space where the embedding coordinates represent the photometric information. Their method fails to provide a mathematical justification for their proposed heat kernel framework or the proposed discretization method. Iglesias and Kimmel [105] used the diffusion distances based on Schrodinger operators incorporating texture data, then compare the histograms of Schrodinger diffusion distances with the earth movers distance. Finally, S. Biasotti et al. [106] proposed the PHOG descriptor as a combination of photometric, hybrid and geometric descriptions into one descriptor for textured 3D object retrieval.

### 5.1.2 Contribution

This chapter develops a mathematical framework for the diffusion geometry on textured shapes. It presents an approach for shape matching and retrieval based on weighted heat kernel signature. As a first contribution, it shows how to include photometric information as a weight over the shape manifold, also propose a novel formulation for heat diffusion over weighted manifolds. As a second contribution this chapter presents a new discretization method for the weighted heat kernel induced by the linear FEM weights. It also proposed a new method to introduce the scale invariance for the weighted heat kernel signature. This proposed scale normalization method eliminates the scale effect with less sensitivity to noise. Finally, the weighted heat kernel signature is used as a shape

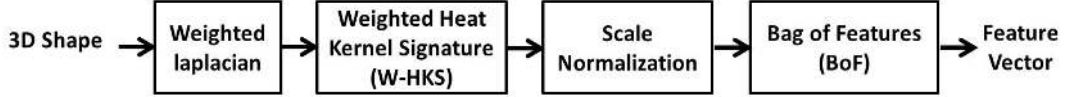


Figure 5.1: Construction of the proposed descriptor.

descriptor. The proposed descriptor encodes both the photometric, and geometric information based on the solution of one equation. Figure 5.1 shows the steps for the construction of the proposed descriptor.

## 5.2 Heat Kernel on Weighted Manifold

The heat diffusion propagation over a manifold  $\mathcal{M}$  is governed by the heat equation (3.13) as described in Section 3.2.2 . A weighted manifold (called also a manifold with density) [107] is a Riemannian manifold  $\mathcal{M}$  endowed with a measure  $\mu$  that has a smooth positive density  $h$  with respect to the Riemannian measure  $\sigma$ . The weighted Laplace operator  $\Delta_{\mathcal{M},\mu}$ , generalizing the Laplace-Beltrami operator, is defined by

$$\begin{aligned}
 \Delta_{\mathcal{M},\mu}U &= \operatorname{div}_{\mu}\nabla U \\
 &= \frac{1}{h}\operatorname{div}(h\nabla U) \\
 &= \frac{1}{h\sqrt{\det g}}\frac{\partial}{\partial x_i}(h\sqrt{\det g})\frac{\partial}{\partial x_i}U
 \end{aligned} \tag{5.1}$$

for any smooth function  $U$  on  $\mathcal{M}$ , where  $g$  be the Riemannian metric on  $\mathcal{M}$ . It is possible to extend  $\Delta_{\mathcal{M},\mu}$  to a self-adjoint operator in  $L^2(\mathcal{M},\mu)$ , which allows one to define the heat semigroup  $e^{-t\Delta_{\mathcal{M},\mu}}$ . The heat semigroup has the integral kernel  $K(\mathbf{x}, \mathbf{y}, t)$ , called the heat kernel of  $(\mathcal{M}, \mu)$ .

For the heat diffusion propagation over a weighted manifold  $(\mathcal{M}, \mu)$  the

weighted heat equation is proposed as:

$$\Delta_{\mathcal{M},\mu}U(\mathbf{x}, t) = -\frac{1}{2}\frac{\partial}{\partial t}U(\mathbf{x}, t) \quad (5.2)$$

with initial condition  $U(\mathbf{x}, 0) = u(\mathbf{x})$  and Dirichlet boundary condition  $U(\mathbf{x}, t) = 0$  for all  $\mathbf{x} \in \partial\mathcal{M}$  and all  $t > 0$ . In the following, a method for the discretization of the weighted heat equation based on the weighted linear finite elements method (FEM) [94] will be described.

### 5.2.1 Finite Element Discretization of the Weighted Heat Equation

The weak formulation of the weighted heat equation (5.2) is obtained by multiplying by the density  $h$  and a test function  $\varphi \in \mathcal{C}^2$  and integrating the resulting relation over the weighted manifold  $(\mathcal{M}, \mu)$

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U(\mathbf{x}, t) h d\sigma + \frac{1}{2} \int_{\mathcal{M}} \varphi \Delta_{\mathcal{M},\mu} U(\mathbf{x}, t) h d\sigma = 0 \quad (5.3)$$

where  $d\sigma$  is the Riemannian measure. Then, using the Green formula it will be:

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U h d\sigma + \frac{1}{2} \int_{\mathcal{M}} \operatorname{div}(h \nabla U) \varphi d\sigma = 0, \quad (5.4)$$

or equivalently,

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U h d\sigma + \frac{1}{2} \int_{\mathcal{M}} (h \nabla U) \cdot \nabla \varphi d\sigma = 0. \quad (5.5)$$

This is called the weak formulation of (5.2). Substituting  $h d\sigma$  by  $d\mu$  then:

$$\int_{\mathcal{M}} \varphi \frac{\partial}{\partial t} U d\mu + \frac{1}{2} \int_{\mathcal{M}} (\nabla U \cdot \nabla \varphi) d\mu = 0 \quad (5.6)$$

Let  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$  be a family of  $n$  linearly independent  $C^2$  functions. An

approximation  $\tilde{U}(\mathbf{x}, t)$  to  $U(\mathbf{x}, t)$  can be defined as:

$$\tilde{U}(\cdot, t) := \sum_{i=1}^n a_i(t) \varphi_i, \quad t > 0. \quad (5.7)$$

Replacing  $U$  by  $\tilde{U}$  in (5.6) then it will be:

$$\sum_{i=1}^n \frac{\partial}{\partial t} a_i(t) \int_{\mathcal{M}} \varphi_i \varphi_j d\mu + \frac{1}{2} \sum_{i=1}^n a_i(t) \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) d\mu = 0, \quad (5.8)$$

or simply

$$\sum_{i=1}^n B(i, j) \frac{\partial}{\partial t} a_i(t) + \frac{1}{2} \sum_{i=1}^n L(i, j) a_i(t) = 0 \quad (5.9)$$

where

$$B(i, j) = \int_{\mathcal{M}} \varphi_i \varphi_j h d\sigma, \text{ and}$$

$$L(i, j) = \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) h d\sigma.$$

To discretize (5.9) let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Choose linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  are equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolates between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $T$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at



vertices  $i, j, k$  respectively, and  $A(T)$  is the area of  $T$  then it can be shown that:

$$\begin{aligned}
\nabla\varphi_i \cdot \nabla\varphi_j &= -\frac{\cot\alpha_k}{2A(T)} \\
|\nabla\varphi|^2 &= \frac{\cot\alpha_j + \cot\alpha_k}{2A(T)} \\
\int_T \varphi_i d\sigma &= \frac{A(T)}{3} \\
\int_T \varphi_i^2 d\sigma &= \frac{A(T)}{6} \\
\int_T \varphi_i \varphi_j d\sigma &= \frac{A(T)}{12} \\
\int_T \varphi_i^3 d\sigma &= \frac{A(T)}{10} \\
\int_T \varphi_i^2 \varphi_j d\sigma &= \frac{A(T)}{30} \\
\int_T \varphi_i \varphi_j \varphi_k d\sigma &= \frac{A(T)}{60}
\end{aligned} \tag{5.10}$$

For the proofs of 5.10 see Appendix D.

Denote by  $h_T$  the average value of a function  $h$  over a triangle  $T$ , and  $N_e(i)$  the set of vertices adjacent to  $i$ . Given  $j \in N_e(i)$  denote by  $T_\alpha$  and  $T_\beta$  the triangles having  $(i, j)$  as an edge and by  $\alpha$  and  $\beta$  the interior angles of  $T_\alpha$  and  $T_\beta$  opposite edge  $(i, j)$ . Then:

$$B(i, j) = (h_i + h_j) \frac{A(T_\alpha) + A(T_\beta)}{60} + \frac{h_{T_\alpha} A(T_\alpha) + h_{T_\beta} A(T_\beta)}{20} \tag{5.11}$$

whenever  $j \in N_e(i)$ ,

$$B(i, i) = \sum_{k \in N_e(i)} A(T_k) \left( \frac{h_i}{15} + \frac{h_{T_k}}{10} \right) \tag{5.12}$$

where  $T_k$  is the counter-clockwise oriented triangle with vertices  $i$  and  $k$ , and  $B(i, j) = 0$  whenever  $i$  and  $j$  are nonadjacent vertices. Likewise,  $L(i, j)$  is

given by

$$L(i, j) = \begin{cases} -\frac{h_{T_\alpha} \cot \alpha + h_{T_\beta} \cot \beta}{2} & \text{if } j \in N_e(i) \\ \sum_{k \in N_e(i)} L(i, k) & \text{if } i = j \end{cases} \quad (5.13)$$

where  $h_{T_\alpha}$ , and  $h_{T_\beta}$  are the average values of a function  $h$  over a triangles  $T_\alpha$ , and  $T_\beta$ . And  $L(i, j) = 0$  otherwise. For the proofs of 5.11, 5.12, and 5.13 see Appendix E.

To compute the solution to Equation 5.2, let us consider the generalized eigensystem  $\{\lambda_i, \phi_i\}_{i=1}^n$  of  $(L, B)$ , which satisfies the relations  $L\phi_i = \lambda_i B\phi_i$ ,  $i = 1, \dots, n$ . Since the Laplacian eigenvectors  $\{\phi_i\}_{i=1}^n$  form a basis of  $\mathbb{R}^n$  and  $(\tilde{U}(\mathbf{x}, t))_{i=1}^n \in \mathbb{R}^n$ , for any  $t \in \mathbb{R}^+$  the solution  $\tilde{U}(\cdot, t)$  can be expressed as  $\tilde{U}(\cdot, t) := \sum_{i=1}^n a_i(t)\phi_i$  where  $\mathbf{a} = (a_i(t))_{i=1}^n$  is the unknown vector.

After solving for the coefficients  $a_i(t)$ , then  $a_i(t) = \exp(-\frac{1}{2}\lambda_i t)\langle u(\mathbf{x}), \phi_i \rangle_B$  where  $u(\mathbf{x})$  is the initial value of  $\tilde{U}(\mathbf{x}, t)$ . Then,

$$\tilde{U}(\cdot, t) := \sum_{i=1}^n \exp(-\frac{1}{2}\lambda_i t)\langle u(\mathbf{x}\phi_i) \rangle_B \phi_i, \quad (5.14)$$

or in a matrix form

$$\tilde{U}(\cdot, t) := \phi D(t) \phi^T B u(\mathbf{x}) \quad (5.15)$$

where  $\phi = [\phi_1, \phi_2, \dots, \phi_n]$ , and

$$D(t) = \text{diag}(\exp(-\frac{1}{2}\lambda_1 t), \exp(-\frac{1}{2}\lambda_2 t), \dots, \exp(-\frac{1}{2}\lambda_n t)).$$

Then the heat kernel will be

$$K(\mathbf{x}, \mathbf{y}, t) := \phi D(t) \phi^T B, \quad (5.16)$$

and the weighted heat kernel signature will be

$$K(\mathbf{x}, \mathbf{x}, t) := B \phi^2 \text{diag}(D(t)) \quad (5.17)$$

The heat kernel signature was introduced by [3, 74, 75] as an intrinsic local shape descriptor based on diffusion scale-space analysis. Here the weighted version is introduced by considering the color information at each vertex as the weight  $h$  as discussed earlier.

### 5.3 Weighted Heat Kernel Signature (W-HKS)

This section, proposes an approach for shape modeling and retrieval using weighted heat kernel signature. Figure 5.1 shows the steps for the construction of the proposed descriptor. The proposed descriptor is based on the BoF representation of the W-HKS calculated all shape vertices at different time scales. This work proposes a novel method to achieve scale-invariance of HK which is shown to be noise-robust in Chapter 4.4. The scale normalization step is applied to the W-HKS before the BoF representation as explained in the following subsection. The proposed descriptor is compact in size, and efficient in computation.

The proposed descriptor is constructed as follows: The W-HKS descriptor is calculated for each triangle mesh based on Equation 5.17. The W-HKS is calculated at all points of a shape over the three normalized color channels (RGB). The W-HKS is calculated at different time scales, A logarithmic scale-space is used with base  $\alpha = 2$  and  $\tau$  ranging from 0.01 to 8 with step 1/16 for each color band. Then the scale normalization step is applied as explained in Section 4.3.2. Considering only the first 20 significant FT components the size of the descriptor now will be  $N \times 20$  for  $N$  vertex shape. Then the Bag of Features (BoF) technique is used to represent the shape as one feature vector. The bags of features were created using the a vocabulary of size 64. Thus the feature vector size is  $64 \times 3$  for the three color channels for any 3D shape.

## 5.4 Experimental Results

The SHREC'13 data set [7] is used to test the performance of the proposed approach. This data set is a collection made of 240 texture shapes, organized in 10 classes, each with 24 models. Each class (humans, four legs animals, vases etc.) contains six null models, that is, two base meshes endowed with three different textures. Each null shape is then modified via four transformations, including two non-metric-preserving deformations, one non-rigid deformation, and one additive Gaussian noise perturbation. All transformations are applied at different strength levels for the ten classes. Also, the same texture may be shared by models in different classes, see Figure 5.2 for sample of different classes for the data set.

The proposed method is compared to the best four methods in the retrieval competition on textured 3D models [7]: a method based on Scale Invariant Heat Kernels combined with the color histogram(A2), the Color-weighted Histograms of Area Projection Transform (Gi), a method based on 2D multi-view and bag-of-features approach (G2), and one method merging a shape description based on geodesic distance matrices with RGB histograms (V2). Also, the proposed method is compared to the PHOG approach proposed in [108].

The average precision-recall curves, Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), and Average Dynamic Recall (ADR) are used as evaluation measures. To compute these measures, two objects are assumed to be belonging to the same class if they share both geometric and texture information. The final scores is the average on all possible queries and it is always less than 1.

Table 5.1 lists the average measures over all classes in the SHREC'13 data set. The table compares the proposed descriptor (W-HKS1) against the (A2) based on Scale Invariant Heat Kernels combined with the color histogram. The NN and ADR measures prove that the proposed descriptor has better performance because it encode the color as well as the geometric information.

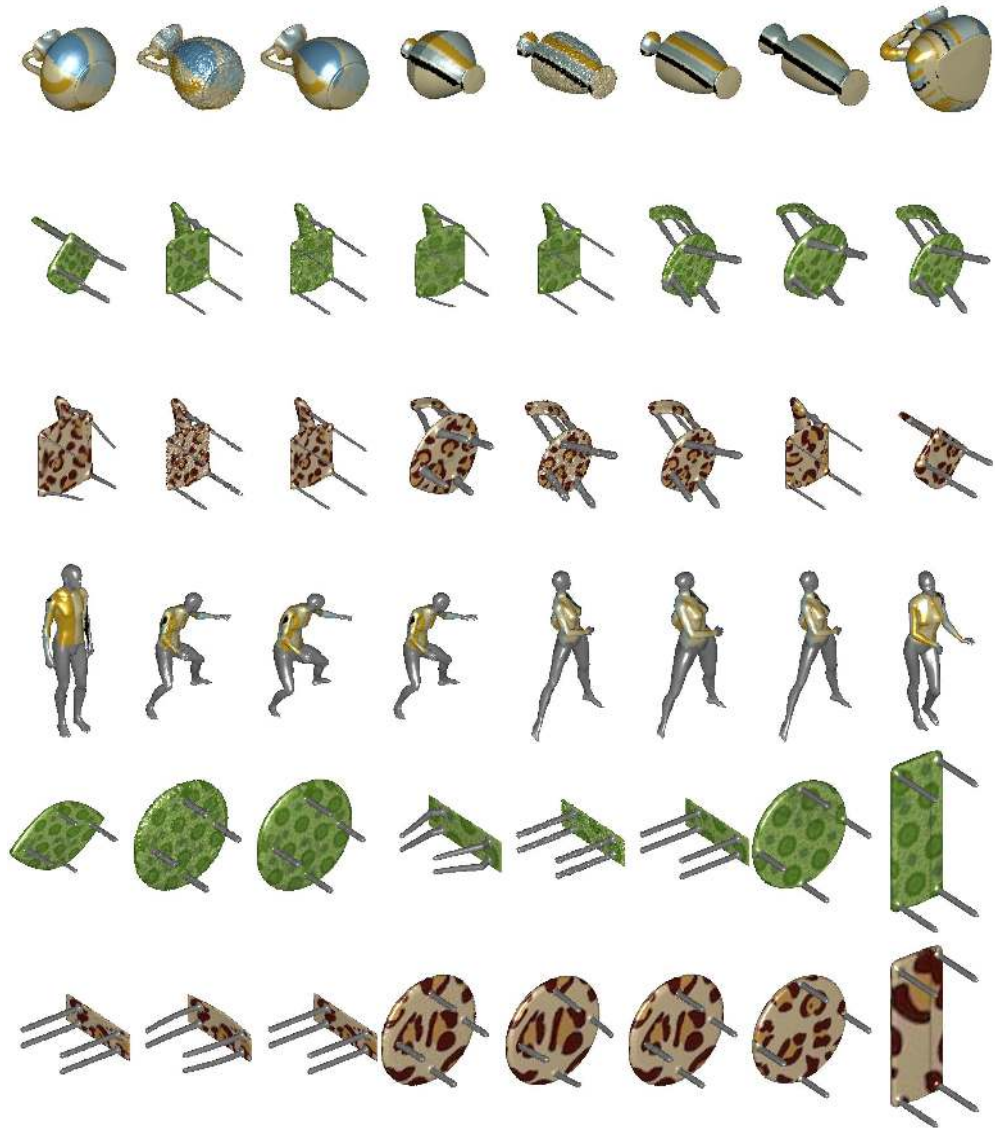


Figure 5.2: Sample of 4 different classes of the SHREC'13 benchmark [7] with different texture and deformation show the challenge of the data set. For each null shape there are a transformed versions of it that include non-rigid deformation, nonmetric-preserving deformations and additive Gaussian noise perturbation.

Run	NN	1-Tier	2-Tier	ADR
A2	0.508	<b>0.561</b>	<b>0.730</b>	0.380
W-HKS1	<b>0.7765</b>	0.5350	0.6458	<b>0.4047</b>

Table 5.1: Results on SHREC'13 data set compares the proposed descriptor (W-HKS1) against the (A2) based on Scale Invariant Heat Kernels combined with the color histogram.

Another version, called (W-HKS2), of the proposed descriptor is formed by appending the color histogram to the original descriptor (W-HKS1). Table 5.2, lists the average measures on the same data set. The table compares the proposed descriptor (W-HKS2) against the five different methods. The highest measures (in bold) clearly show that the proposed descriptor has the best performance.

Figure 5.3 shows some retrieval results. The models are ordered from left to right. The first column represents the query model. The retrieved objects per each query are ranked from left to right based on the distance measure (L1-Norm) between the query and all shapes in the data set. Only the first 15 retrieved models are shown. The figure clearly demonstrates the high retrieval rate of the approach. It can be observed that the proposed method has retrieved the similar shapes with similar texture first.

Figure 5.4 shows the performances of all methods in terms of average precision-recall curves. The larger the area below such a curve, the better the performance under examination. This figure shows that the proposed descriptor indeed achieves the highest performance over the other five methods.

## 5.5 Summary

This chapter has addressed the problem of textured 3D shapes representation. A new approach for shape matching and retrieval based on Weighted Heat Kernel Signature (W-HKS) is presented. The color information is used as a weight over the shape manifold. A novel formulation for heat diffusion over



Figure 5.3: Shape retrieval results of SHREC'13 data set. Left: queries. Right: First 15 matches using the CHKS descriptor.

Run	NN	1-Tier	2-Tier	ADR
A2	0.508	0.561	0.730	0.380
G1	0.788	0.658	0.748	0.470
G2	0.898	0.733	0.893	0.508
V2	0.879	0.764	0.904	0.520
PHOG	<b>0.951</b>	0.773	0.899	0.534
W-HKS2	0.9242	<b>0.8338</b>	<b>0.9257</b>	<b>0.5701</b>

Table 5.2: Results on SHREC'13 data set lists the average measures on SHREC'13 data set. The table compares the proposed descriptor (W-HKS2) against five different methods.

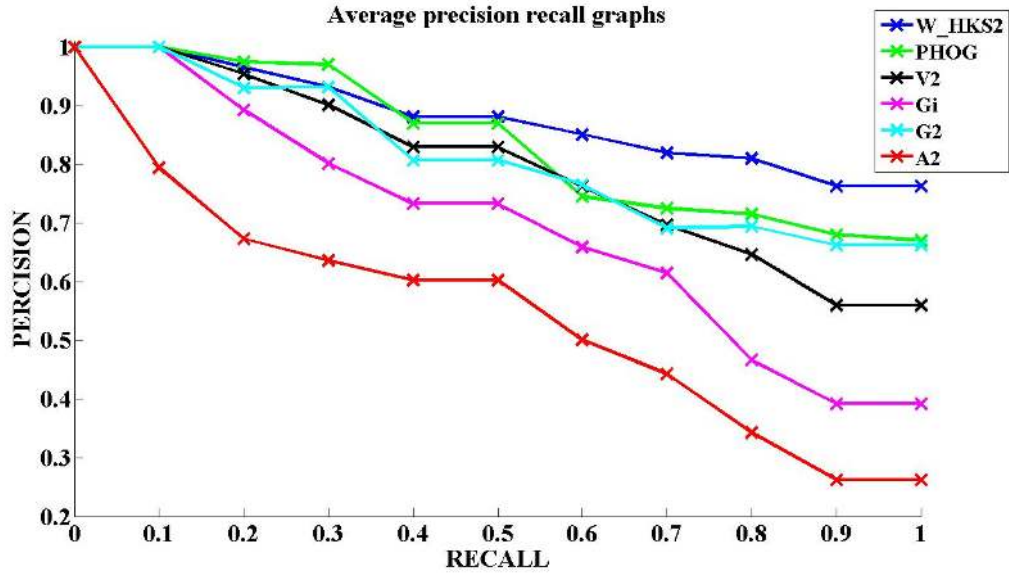


Figure 5.4: Shape retrieval results of SHREC'13 data set. Precision-recall curves for all methods.

weighted manifolds is proposed. Also this chapter presented a new discretization method for the weighted heat kernel based on FEM. It proposed a new method to introduce the scale invariance for the weighted heat kernel signature. The "bag of features" (BoF) approach is used to construct compact and informative shape descriptors. Finally, the weighted heat kernel signature is used as a shape descriptor.

The experimental results have shown that the proposed descriptor can achieve high performance on SHREC'13 benchmark data sets. The proposed approach has outperformed state-of-the-art approaches (five different methods) for textured shapes representation and retrieval. Different evaluation measures approved the high accuracy of the proposed framework.



## CHAPTER 6

# HEAT FRONT PROPAGATION CONTOURS FOR 3D FACE RECOGNITION

Face recognition is a key biometric method aiming at identifying individuals by the features of face. Due to the challenges facing face recognition from 2D images, researchers have resorted to 3D face recognition. The work in this chapter is motivated by the recent and remarkable success of heat-based features for 3D object classification and retrieval. This chapter proposes an approach for 3D face recognition based on the front contours of heat propagation over the face surface. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The propagation contours are used to successfully discriminate the various faces. The proposed approach is evaluated on the largest publicly available database of 3D facial images and successfully compared to the state-of-the-art approaches in the literature.

### 6.1 Introduction

Face recognition is one of the biometric methods identifying individuals by the features of face. Automatic face recognition has evolved from small scale research systems to a wide range of commercial products. Therefore, computer vision methodology for automatic face recognition has become an attractive research area in the past three decades (for more details see [109, 110]).

In the beginning, most efforts were directed towards 2D facial recogni-

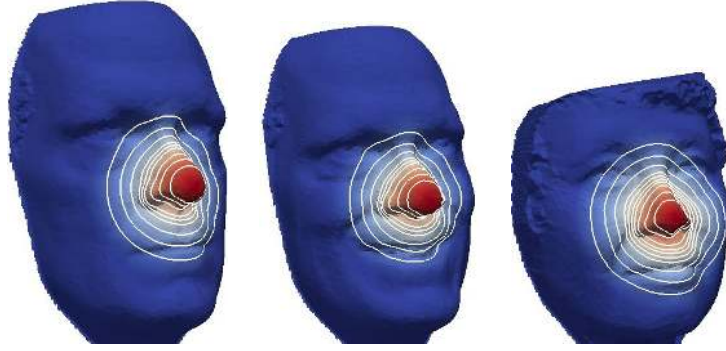


Figure 6.1: The basic idea of the proposed approach: The heat propagation front contours for a virtual heat source at some facial landmarks (nose tip here) are used as discriminating features between different faces. First and second faces are for the same subject thus having similar contours, while the third face has different contours because it is for another subject.

tion [111]. However, there are challenging issues, when using 2D images for face recognition. The face needs to be well lit by controlled light sources in order to acquire high quality images. In addition, faces can be obstructed by hair, glasses, hats, scarves etc. Variances, such as disguises, makeup, change over time (aging), expression and pose, might have a minor to major negative impact on a face recognition system, decreasing its ability to recognize faces. A study of commercial face recognition systems called the Face Recognition Vendor Test (FRVT) 2002-2006 [112] concluded that identification and verification accuracy suffers significantly when there were differences in pose and illumination between gallery enrollment and probing.

To address pose and illumination challenges, researchers have resorted to 3D face recognition (for more details see [113, 114, 115]). Facial 3D geometry either can be acquired using 3D sensing devices such as laser scanners [116] or reconstructed from one or more images [117]. Using 3D sensing devices has proven to be effective in 3D face recognition [118]. Also a lot of work has been done on 3D face reconstruction either from stereo images or from a single image, which enables the extraction of 3D information from 2D acquired facial images. There is a wide range of approaches for 3D face reconstruction including shape from shading [117],

space carving [119], shape from motion [120], stereo reconstruction, shape from texture and analysis by synthesis using morphable models [121]. Each approach has its own assumptions and might work in certain cases and might not work in other cases when the assumptions are violated.

Face recognition from 3D has some advantages over 2D facial images. Their pose can be easily corrected by rigid rotations in 3D space. They also provide structural information about the face (e.g., surface curvature and geodesic distances), which cannot be obtained from a single 2D image. Lastly, 3D face recognition algorithms have been shown to be robust to variations in illumination conditions during image acquisition.

3D face recognition approaches can be divided to three main categories. The first category of Local features approaches utilizes local features, such as SIFT for meshes and face symmetry [122]. In the second category, deformable template-based approaches have been proposed. As an example, Kakadiaris et al. [123] utilize an annotated face model to study geometrical variability across faces. The annotated face model is deformed elastically to fit each face, thus matching different anatomical areas such as the nose, eyes and mouth. In the third category of surface-distance based approaches, distances between feature points on the face surface are employed. Gupta et al. [8] use Euclidean/geodesic distances between fiducial points, in conjunction with linear classifiers. As stated earlier, the problem of automated detection of fiducial points is non-trivial and hinders automation of these methods. Also Bronstein et al. [124] provide a limited experimental illustration of this invariance by comparing changes in surface distances with the Euclidean distances between corresponding points on a canonical face surface. To handle the open mouth problem, they first detect and remove the lip region, and then compute the surface distance in presence of a hole corresponding to the removed part.

This method is motivated by the recent and remarkable success of heat-

based features for 3D object classification and retrieval [3, 31]. The idea is to propose an approach for 3D face recognition based on the front contours of heat propagation over the face surface, see Figure 6.1. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The extraction of those fiducial landmarks is fully automated. The proposed approach encodes the local face features as well as the diffusion distance over the surface around these landmarks. The propagation contours are used to successfully discriminate the various faces. The proposed approach is evaluated on the Texas 3D face recognition database [8] as it is the largest publicly available database of 3D facial images acquired using a stereo imaging system. It is also compared to the state-of-the-art approaches in the 3D face recognition literature.

## 6.2 Proposed Approach for 3D Face Recognition

This work presents an approach for 3D face recognition based on heat kernel (HK). It develops a new approach to extract the geometric points that have the same diffusion distance from a source landmark point. These points lie on a contour around a source landmark point. These contours are unique for each face, and using it as a face descriptor is able to discriminate between human faces. This section will describe how the landmarks points are detected. Then it will give the details about the solution of heat equation over a manifold, and the reconstruction of the heat kernel based on the eigenfunctions and eigenvalues of the Laplace-Beltrami operator. This will allow to explain how the contours are extracted around the facial landmark. Figure 6.2 shows the Proposed approach framework.

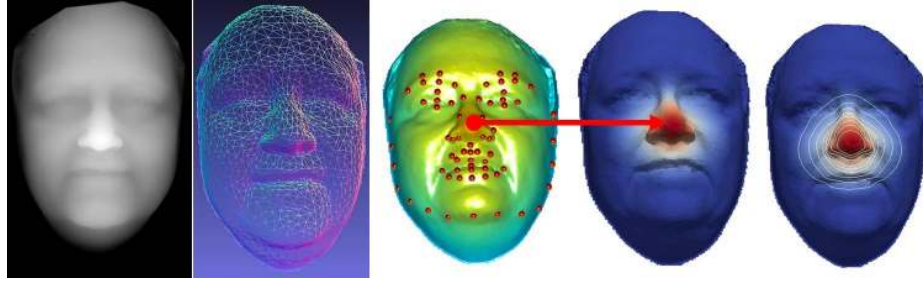


Figure 6.2: Proposed approach: Given a depth image, triangulation to get 3D mesh, feature point extraction, heat kernel reconstruction for each feature point, 3D contours extraction, then contour sampling.

### 6.2.1 Facial Landmark Detection

The relationship among facial feature positions is commonly modeled as a single Gaussian distribution function [125], which is the model used by the Active Appearance Model (AAM) and Active Shape Model (ASM) algorithms. Everingham et.al [126] modelled the probability distribution over the joint position of the features using a mixture of Gaussian trees. The appearance of each facial feature is assumed independent of the other features and is modelled using a variation of the AdaBoost algorithm [127]. Figure 6.3 shows the facial features detected on sample faces.

This work uses the method proposed in [126] to detect a small set of facial landmarks (exactly nine points). The landmarks are extracted from the 2D image and then mapped to the 3D surface. These points are used to initialize the Extended Active Shape Model (STASM) [128] for mesh fitting, where the STASM base mesh is warped to these nine points. The output of this step is 68 facial landmarks. Figure 6.3 illustrates an example for fitting STASM-based meshes on sample faces of Texas 3D data set. Later it will be shown that just a subset of 12 points will be used.

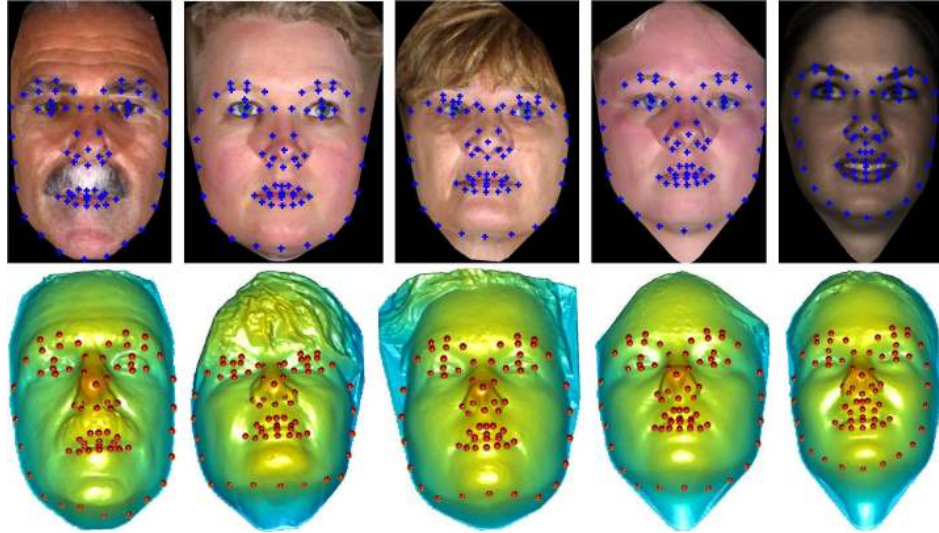


Figure 6.3: Facial landmarks detected on sample faces from Texas 3D faces data set.

## 6.2.2 Heat Front Propagation Contours

Section 3.4.2 provided a solution of the heat equation based on a linear Finite Element Method (FEM) approximation to derive a discrete heat kernel. To compute the solution to (3.13) the weak formulation through the Galerkin formulation is used, as described in Section 3.4.2, then the heat kernel will be:

$$K(\mathbf{x}, \mathbf{y}, t) := \phi D(t) \phi^T B, \quad (6.1)$$

The heat kernels are calculated at each detected landmark point for different time scales. Then a set of 3D contours on the face surface are extracted based on the heat kernel. These contours can be used to discriminate between the different faces.

Figure 6.4 shows the first four eigenfunctions of the Laplace-Beltrami operator of a human face. The first  $n = 100$  eigenfunctions and eigenvalues of the Laplace-Beltrami operator will be used to construct the heat kernels in the following subsection.

Figure 6.5 shows the HKs induced from three different critical point for

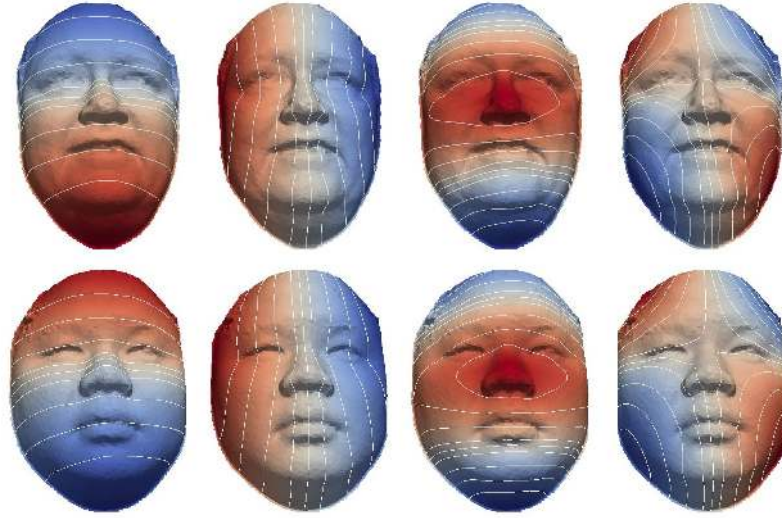


Figure 6.4: The first four non-zero eigenfunctions of the Laplace-Beltrami operator of human faces

one sample face at different time samples. For small time, the signature captures local shape information. As time elapses, the signature tends to capture more global shape details.

### 6.2.3 Contours Matching

After calculating the heat kernels at each point, the 3D contours are extracted at the 3D point on the face surface that have equal heat values. A predetermined number of the contours is used around each point. Then each contour is sampled with a fixed number of points. This representation gives a finite and ordered set of 3D points per face. To match two faces, the Iterative Closest Point (ICP) algorithm is used to estimate rigid transformation parameters between the corresponding point sets for the found contours on the two faces. The  $L2 - norm$  distance between the contour points of the probe face and gallery faces after registration is used as the distance measure, and the gallery faces are ranked based on this distance measure. The Iterative Closest Point (ICP) algorithm is described in Appendix F.

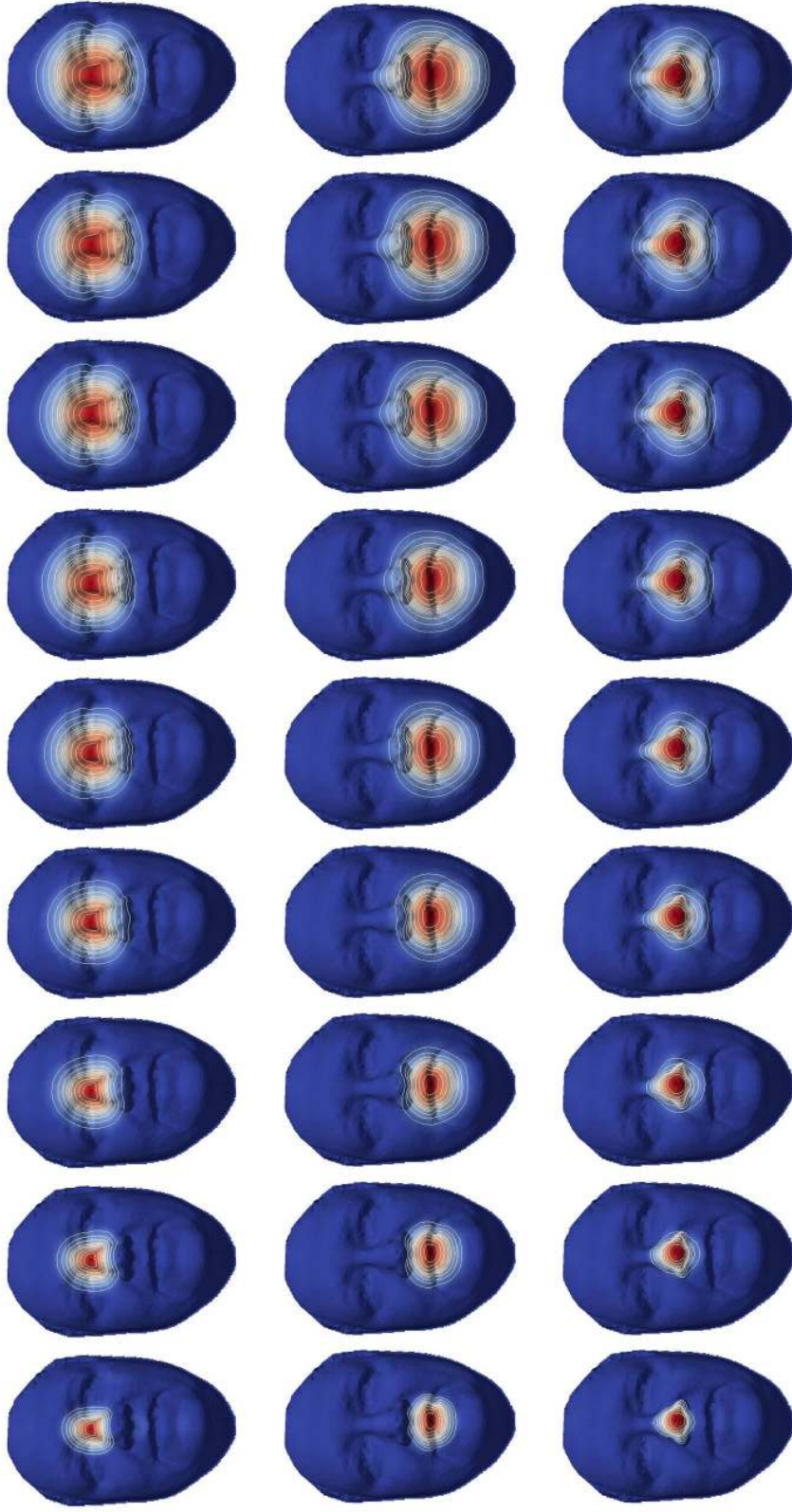


Figure 6.5: HK induced from three different landmarks point detected on the face shape at seven different time samples. Colors represent the values of the heat induced from the source point to each vertex of the shape. For small time, the HK is more local. As time progresses, the HK becomes more global.



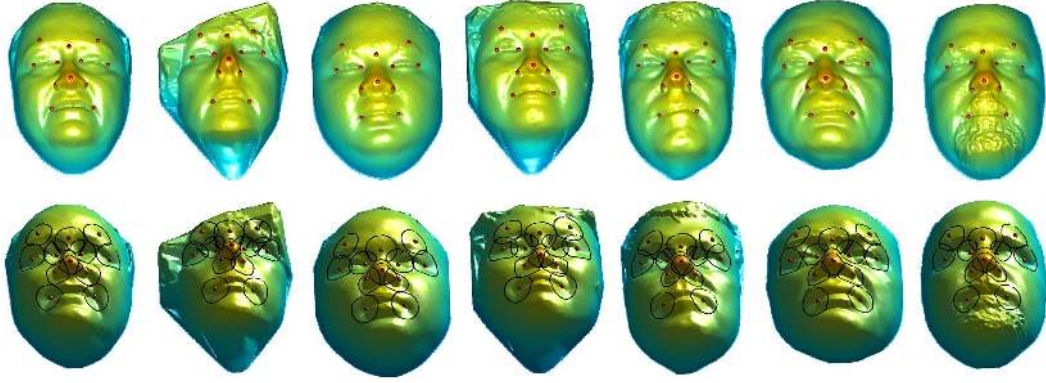


Figure 6.6: Upper row: The selected 12 landmarks, Lower row: The reconstructed one contour around each point. These 12 contours are used as a descriptor for the 3D face.

### 6.3 Experimental Results

To assess the performance of the proposed approach, a test is done on the Texas 3D face recognition database [8, 129]. Currently, Texas 3D face recognition database is the largest publicly available database of 3D facial images acquired using a stereo imaging system. The database contains 1149 3D models of 118 adult human subjects. The number of images of each subject varies from 1 per subject to 89 per subject. The subjects ages range from  $\sim 22 - 75$  years. The database includes images of both males and females from the major ethnic groups of Caucasians, Africans, Asians, East Indians, and Hispanics. Samples of the RGB image, depth image and the 3D mesh are shown in Figure 6.7. The 3D mesh are created by triangulating the depth image. This experiment uses a gallery set of 105 subjects (only one 3D model for each subject is used). The probe set is selected from 480 sessions (3D faces).

A series of initial experiments are done to test different number of contours around each facial landmark. It was found that using as few as one contour can provide a good recognition rate. Using more contours could improve the results further but at the cost of increased computational time. In the reported results, one contour per each landmark is employed. Also only 12 landmark points are

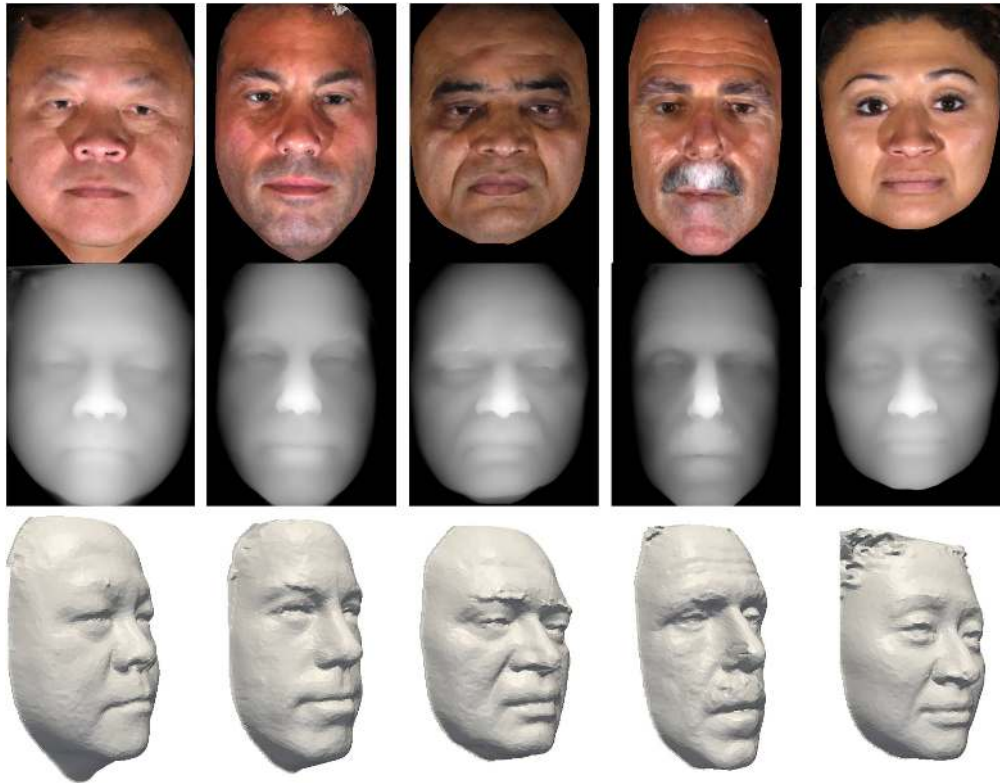


Figure 6.7: Samples of Texas 3D faces data set [8], from top to bottom The RGB image, the depth image and the 3D mesh.

used in this experiment, Figure 6.6 shows sample of 3D faces with the used feature points and the extracted contours.

For the sake of comparison, the results of five methods are shown: the eigensurfaces of Chang et al. [130], fishersurfaces of BenAbdelkader and Griffin [131], and ICP algorithms Lu et al. [132], the anthroface 3D algorithm [8], based on 25 manually located points, and the anthroface 3D algorithm that employed 10 automatically located points [8]. using the same probe and gallery sets.

Figure 6.9 shows the Cumulative Rank Curves (CMC) for the five algorithms together with the proposed approach results. The proposed approach has achieved a 100% recognition rate at rank 4. The results indeed confirm the superior performance of the proposed approach over the state-of-the-art approaches for 3D face recognition.

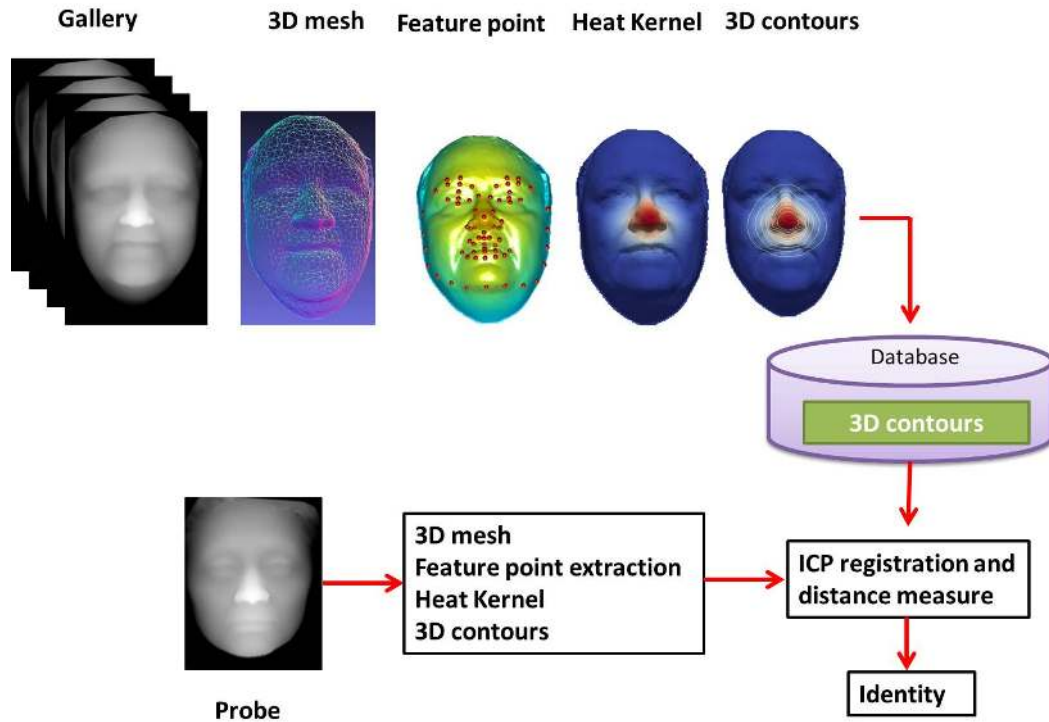


Figure 6.8: 3D face recognition framework.

## 6.4 Summary

This chapter, addressed the problem of 3D face recognition. It presented a new approach based on the front contours of heat propagation over the face surface. This was motivated by the recent and remarkable success of heat-based features (Chapter 4.4) for 3D object classification and retrieval [31]. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The extraction of those fiducial landmarks is fully automated. The approach encodes the local face features as well as the diffusion distance over the surface around these landmarks.

The proposed approach has been evaluated on the Texas 3D Face Recognition Database [8] as it is the largest publicly available database of 3D facial images acquired using a stereo imaging system. The results have demonstrated the superior performance of the proposed approach over several state-of-the-art

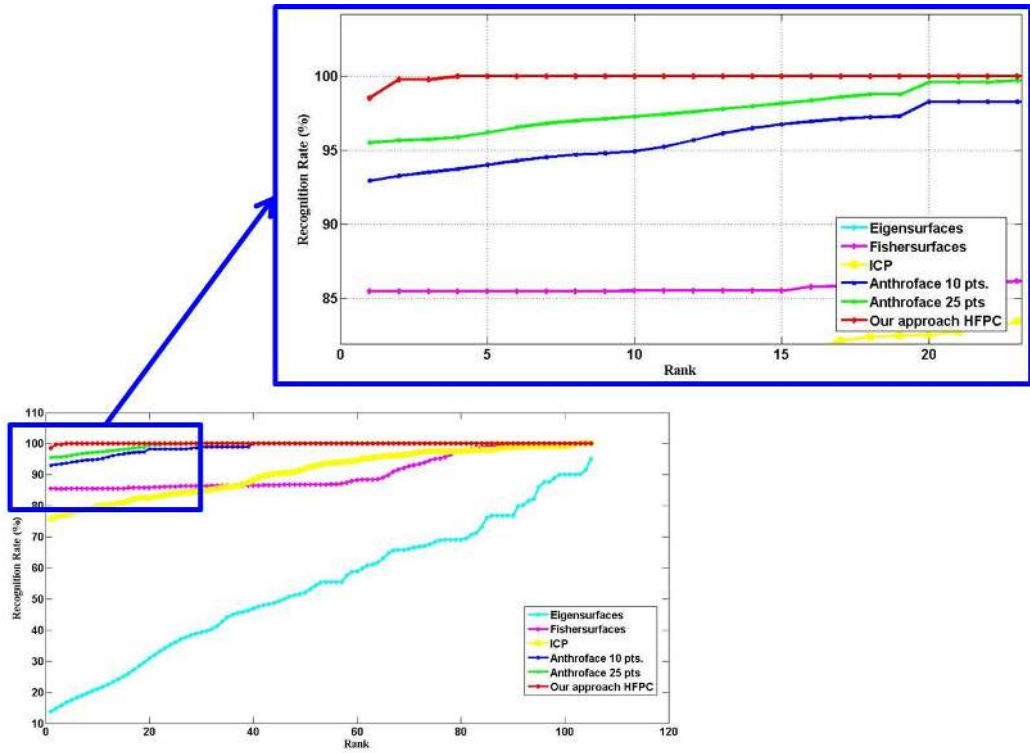


Figure 6.9: The CMC curves, for the eigensurfaces, fishersurfaces, ICP algorithms, anthroface 3D algorithm based on 25 manually located points, the anthroface 3D algorithm that employed 10 automatically located points [8], and the proposed approach HFPC.

approaches in the 3D face recognition literature. Future work may include testing the proposed approach on different databases.

## CHAPTER 7

### CONCLUSIONS AND FUTURE WORK

This dissertation has proposed a number of advances in modeling 3D non-rigid shapes. The findings from this dissertation make several contributions to the current literature which can be summarized as follows:

Chapter 4, proposed a new method for feature detection on 3D shapes. It proposed to use the first non-trivial Laplace-Beltrami eigenfunction to detect a small number of sparse critical points on the surface of the shape. An observation shows that shapes belonging to the same class consistently have almost the same number of critical points, whereas these numbers differ from one class to another. As such, this number can be used as one of the discriminatory features between the different classes.

As a second contribution, two new feature descriptors are proposed, one based on the heat kernels, and another constructed from scale-invariant heat kernels. For the first approach, a feature vector (descriptor) computed at these critical points is sufficient for shape matching. Sparse representation of the descriptor can be obtained to reduce its dimensionality before shape classification. The proposed descriptors are compact in size, the feature vector can be down to 64, and efficient in computation since it is constructed at sparse points.

The experimental results have shown that the proposed descriptor can achieve state-of-the-art results on the two benchmark data sets. The results have indeed confirmed the improved performance of the sparse representation over the original non-reduced descriptors, yet reducing the time and space complicity for the shape retrieval problem. In addition, results have shown that using one critical

point will not sacrifice the performance.

The second approach presented a new approach for shape matching and retrieval based on scale-invariant heat kernel (HK). The proposed descriptor is based on the histograms of HKs at different time scales computed at these critical points. It thus can capture the local (at small time scales) as well as global shape information (at bigger time scales). A new local transformation method to introduce scale-invariance has been also proposed, which is shown to be simpler, more computationally efficient and noise-robust compared to a well-known method. The collaborative classification is used as a feature classifier

The experimental results have shown that the proposed descriptor can achieve high performance on two popular benchmark data sets. The proposed approach has outperformed state-of-the-art approaches for shape representation and retrieval. The results have demonstrated that the proposed descriptor enjoys many of the properties desirable in a shape descriptor: it is isometry-invariant, able to deal with the non-rigid shape deformation. It is independent of the object's size, stable to small changes in the shape, and less sensitive to noise. As the descriptor is computed at a number of points on the shape, rather than the whole shape, it is computationally efficient. It has similarity, uniqueness and scope properties; it has retrieved almost all the database shapes truly matching the same query shape. An important observation from the experiments is that the proposed approach is more able to handle data under several distortion scenarios (white noise, shot-noise, and scale changes) than the well-known Shape Google and Shape-DNA approaches.

Chapter 5, addressed the problem of textured 3D shapes representation. It presented a new approach for shape matching and retrieval based on Weighted Heat Kernel Signature (W-HKS). This chapter proposed to use the color information as a weight over the shape manifold. It also proposed a novel formulation for heat diffusion over weighted manifolds. Then it presented a new discretization

method for the weighted heat kernel based on FEM. This chapter also proposed a new method to introduce the scale invariance for the weighted heat kernel signature. The "bag of features" (BoF) approach is used to construct compact and informative shape descriptors. Finally, the weighted heat kernel signature is used as a shape descriptor. The experimental results have shown that the proposed descriptor can achieve high performance on SHREC13 benchmark data sets. The proposed approach has outperformed state-of-the-art approaches (five different methods) for textured shapes representation and retrieval. Different evaluation measures approved the high accuracy of the proposed framework.

Chapter 6, addressed the problem of 3D face recognition, in the framework of the BOSS project in the CVIP lab. It presented a new approach based on the front contours of heat propagation over the face surface. The front contours are extracted automatically as heat is propagating starting from a detected set of landmarks. The extraction of those fiducial landmarks is fully automated. The proposed approach encodes the local face features as well as the diffusion distance over the surface around these landmarks. The proposed approach has been evaluated on the largest publicly available database of 3D facial images acquired using a stereo imaging system. The results have demonstrated the superior performance of the proposed approach over several state-of-the-art approaches in the 3D face recognition literature.

The analysis also lead to a number of interesting observations, some of which may be considered as lines of future research. Below one outlines some limitations as well as several important directions of further research.

- Using the proposed approach to address dense correspondence between non-rigid shapes.
- Use the proposed approach with the properties of the Laplace-Beltrami eigenfunction for 3D mesh segmentation.

- Use the proposed approach for view point selection for 3D object by selecting the most informative views that collectively provide the most descriptive presentation of the surface.
- Address the problem of joint diagonalization for the construction of coupled bases of the Laplacians of multiple shapes.

In conclusion, this dissertation has presented an analytical framework for modeling 3d non-rigid shapes, textured 3d non-rigid shapes, and 3D face recognition. One believes that there is still much progress to be made in the future.



## REFERENCES

- [1] Yi Fang, Mengtian Sun, and Karthik Ramani. Temperature distribution descriptor for robust 3d shape retrieval. In *NORDIA'11*, 2011.
- [2] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace-beltrami spectra as "shape-dna" of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.
- [3] Alexander M. Bronstein, Michael M. Bronstein, Leonidas J. Guibas, and Maks Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph.*, 30(1):1, 2011.
- [4] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Efficient computation of isometry-invariant distances between surfaces. *SIAM J. Scientific Computing*, 28(5):1812–1836, 2006.
- [5] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Calculus of nonrigid surfaces for geometry and texture manipulation. *IEEE Trans. Vis. Comput. Graph.*, 13(5):902–913, 2007.
- [6] Martin Reuter. Hierarchical shape segmentation and registration via topological features of laplace-beltrami eigenfunctions. *International Journal of Computer Vision*, 89(2):287–308.
- [7] Andrea Cerri, Silvia Biasotti, Mostafa Abdelrahman, Jesús Angulo, K. Berger, Louis Chevallier, Moumen T. El-Melegy, Aly A. Farag, F. Lefebvre, Andrea Giachetti, H. Guermoud, Y.-J. Liu, Santiago Velasco-Forero,

- Jean-Ronan Vigouroux, C.-X. Xu, and J.-B. Zhang. Shrec'13 track: Retrieval on textured 3d models. In *3DOR*, pages 73–80, 2013.
- [8] Shalini Gupta, Mia K. Markey, and Alan C. Bovik. Anthropometric 3d face recognition. *International Journal of Computer Vision*, 90(3):331–349, 2010.
- [9] Panagiotis Papadakis, Ioannis Pratikakis, Stavros J. Perantonis, and Theoharis Theoharis. Efficient 3d shape matching and retrieval using a concrete radialized spherical projection representation. *Pattern Recognition*, 40(9):2437–2452, 2007.
- [10] Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19, 2004.
- [11] Johan W. H. Tangelder and Remco C. Veltkamp. A survey of content based 3d shape retrieval methods. *Multimedia Tools Appl.*, 39(3):441–471, 2008.
- [12] Johan W. H. Tangelder and Remco C. Veltkamp. A survey of content based 3d shape retrieval methods. In *SMI*, pages 145–156, 2004.
- [13] Ran Gal, Ariel Shamir, and Daniel Cohen-Or. Pose-oblivious shape signature. *IEEE Trans. Vis. Comput. Graph.*, 13(2):261–271, 2007.
- [14] Mohamed S. Hassouna. *Variational Curve Skeletons: Novel Framework For Structural Shape Representation Of Volumetric Objects*. PhD thesis, ECE, University of Louisville, Louisville, KY, 2006.
- [15] Petros Daras and Apostolos Axenopoulos. A 3d shape retrieval framework supporting multimodal queries. *International Journal of Computer Vision*, 89(2-3):229–247, 2010.
- [16] Philip Shilane, Patrick Min, Michael M. Kazhdan, and Thomas A.

- Funkhouser. The princeton shape benchmark. In *SMI*, pages 167–178, 2004.
- [17] C. Brechbuhler, G. Gerig, and O. Kubler. Parametrization of closed surfaces for 3-d shape description. *Journal of Computer Vision and Image Understanding*, 61:154–170, 1995.
- [18] Natraj Iyer, Subramaniam Jayanti, Kuiyang Lou, Yagnanarayanan Kalyanaraman, and Karthik Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, 37(5):509–530, 2005.
- [19] Andrew Edie Johnson and Martial Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(5):433–449, 1999.
- [20] Sameh M. Yamany and Aly A. Farag. Surface signatures: An orientation independent free-form surface representation scheme for the purpose of objects registration and matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1105–1120, August 2002.
- [21] Michael Kolomenkin, Ilan Shimshoni, and Ayellet Tal. On edge detection on surfaces. In *CVPR*, pages 2767–2774, 2009.
- [22] Roberto Toldo, Umberto Castellani, and Andrea Fusiello. Visual vocabulary signature for 3d object retrieval and partial matching. In *3DOR*, pages 21–28, 2009.
- [23] Mirela Ben-Chen and Craig Gotsman. Characterizing shape using conformal factors. In *3DOR*, pages 1–8, 2008.
- [24] Asi Elad (Elbaz) and Ron Kimmel. Bending invariant representations for surfaces. In *CVPR (1)*, pages 168–174, 2001.

- [25] Haibin Ling and David W. Jacobs. Using the inner-distance for classification of articulated shapes. In *CVPR (2)*, pages 719–726, 2005.
- [26] Alexander M. Bronstein, Michael M. Bronstein, Alfred M. Bruckstein, and Ron Kimmel. Analysis of two-dimensional non-rigid shapes. *International Journal of Computer Vision*, 78(1):67–88, 2008.
- [27] Martin Reuter, Franz-Erich Wolter, Martha Elizabeth Shenton, and Marc Niethammer. Laplace-beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. *Computer-Aided Design*, 41(10):739–755, 2009.
- [28] Raif M. Rustomov. Laplace-beltrami eigenfunctions for deformation invariant shape representation. In *Symposium on Geometry Processing*, pages 225–233, 2007.
- [29] A. Kovnatsky, D. Raviv, M. M. Bronstein, A. M. Bronstein, and R. Kimmel. Geometric and photometric data fusion in non-rigid shape analysis. *Numerical Mathematics: Theory, Methods and Applications (NM-TMA)*, 6(1):199–222, 2013.
- [30] Mostafa Abdelrahman, Moumen T. El-Melegy, and Aly A. Farag. Heat kernels for non-rigid shape retrieval: Sparse representation and efficient classification. In *CRV*, pages 153–160, 2012.
- [31] Mostafa Abdelrahman, Moumen T. El-Melegy, and Aly A. Farag. 3d object classification using scale invariant heat kernels with collaborative classification. In *ECCV Workshops (1)*, pages 22–31, 2012.
- [32] R. C. Veltkamp and M. Hagedoorn. State of the art in shape matching. *Principles of visual information retrieval*, 87,, 2001.

- [33] J. W. H. Tangelder and R. C. Veltkamp. A survey of content based 3d shape retrieval methods. *Multimedia Tools and Applications*, 39, 3,:441471, 2008.
- [34] Cordelia Schmid, Roger Mohr, and Christian Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.
- [35] Eric Grimson and Thomas Lozano-Perez. Localizing overlapping parts by searching the interpretation tree. In *IEEE Trans. on Pattern Analysis and Machine Intelligence*, page 469482, 1987.
- [36] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [37] Ronen Basri and David.W. Jacobs. Recognition using region correspondences. In *International Journal of Computer Vision*, page 141162, 1996.
- [38] Y. Zhang, M. Brady, , and S. Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, 2001.
- [39] G. Bouchard and B. Triggs. Hierarchical part-based visual object categorization. In *CVPR*, June 2005.
- [40] M.Fussenegger A. Opelt, A. Pinz and P.Auer. Generic object recognition with boosting. In *PAMI*, March 2006.
- [41] A. Awan S. Agarwal and D. Roth. Learning to detect objects in images via a sparse, part-based representation. In *PAMI*, page 26(11):14751490, November 2004.

- [42] A. Leonardis B. Leibe and B. Schiele. Combined object categorization and segmentation with an implicit shape model. In *In ECCV Workshop on Statistical Learning in Computer Vision*, page 1732, Prague, Czech Republic, May 2004.
- [43] T. L. Berg A. C. Berg and J. Malik. hape matching and object recognition using low distortion correspondence. In *CVPR*, June 2005.
- [44] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *n International Conference on Computer Vision*, page 14701477, Oct. 2003.
- [45] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Proceedings of the Fourth Alvey Vision Conference*, pages 147–152, Manchester University, August-September 1988. The University of Sheffield Printing Unit.
- [46] S. M. Smith and J. M. Brady. Susan - a new approach to low level image processing. *International Journal of Computer Vision*, 23:45–78, 1995.
- [47] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *Proceedings of the 7th European Conference on Computer Vision-Part I (ECCV '02)*, pages 128–142, Copenhagen, Denmark, 2002. Springer-Verlag.
- [48] Tony Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch a method for focus-of-attention. *nternational Journal of Computer Vision*, 11(3):283318, 1993.
- [49] Krystian Mikolajczyk and Cordelia Schmid. Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1):63–86, 2004.

- [50] David G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 1150–1157, Corfu, Greece, 1999.
- [51] J.M. Morel and G. Yu. Asift: A new framework for fully affine invariant image comparison. In *SIAM Journal on Imaging Sciences*, page 438469, 2009.
- [52] Alaa E. Abdel-Hakim and Aly A. Farag. CSIFT: A SIFT descriptor with color invariant characteristics. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR06)*, pages 1978–1983, New York City, NY, 17-22 June 2006.
- [53] W. Cheung and G. Hamarneh. N-sift: N-dimensional scale invariant feature transform for matching medical images. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pages 720–723, 2007.
- [54] T. Kadir and M. Brady. Scale, saliency and image description. In *International Journal of Computer Vision*, page 83105, 2001.
- [55] M. Urban J. Matas, O. Chum and T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *n Proceedings of the British Machine Vision Conference*, page 384393, 2002.
- [56] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape context: A new descriptor for shape matching and object recognition. In Todd K. Leen, Thomas G. Dietterich, and Volker Tresp, editors, *Advances in Neural Information Processing Systems 13*, pages 831–837. MIT Press, 2001.
- [57] Tinne Tuytelaars Luc Van Gool Herbert Bay, Andreas Ess. Surf: Speeded up robust features. In *Computer Vision and Image Understanding (CVIU)*., pages Vol. 110, No. 3, pp. 346–359, 2008.

- [58] E. Tola, V. Lepetit, and P. Fua. DAISY: An Efficient Dense Descriptor Applied to Wide Baseline Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):815–830, May 2010.
- [59] Timo Ojala, Matti Pietikainen, and David Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*, volume 1, pages 582–585. IEEE, 1994.
- [60] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *ECCV (4)*, pages 778–792, 2010.
- [61] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. Orb: An efficient alternative to sift or surf. In *ICCV*, pages 2564–2571, 2011.
- [62] Stefan Leutenegger, Margarita Chli, and Roland Siegwart. Brisk: Binary robust invariant scalable keypoints. In *ICCV*, pages 2548–2555, 2011.
- [63] Zhenhua Wang, Bin Fan, and Fuchao Wu. Local intensity order pattern for feature description. In *ICCV*, pages 603–610, 2011.
- [64] Zhouhui Lian, Paul L. Rosin, and Xianfang Sun. Rectilinearity of 3d meshes. *International Journal of Computer Vision*, 89(2-3):130–151, 2010.
- [65] Mauro R. Ruggeri, Giuseppe Patanè, Michela Spagnuolo, and Dietmar Saupe. Spectral-driven isometry-invariant matching of 3d shapes. *International Journal of Computer Vision*, 89(2-3):248–265, 2010.
- [66] Philip Shilane and Thomas A. Funkhouser. Selecting distinctive 3d shape descriptors for similarity retrieval. In *SMI*, page 18, 2006.



- [67] Thomas A. Funkhouser, Patrick Min, Michael M. Kazhdan, Joyce Chen, J. Alex Halderman, David P. Dobkin, and David Pokrass Jacobs. A search engine for 3d models. *ACM Trans. Graph.*, 22(1):83–105, 2003.
- [68] Michael M. Kazhdan, Thomas A. Funkhouser, and Szymon Rusinkiewicz. Symmetry descriptors and 3d shape matching. In *Symposium on Geometry Processing*, pages 117–126, 2004.
- [69] Michael Elad, Ayelet Tal, and Sigal Ar. Content based retrieval of vrml objects - an iterative and interactive approach. pages 97–108, 2001.
- [70] Robert Osada, Thomas A. Funkhouser, Bernard Chazelle, and David P. Dobkin. Shape distributions. *ACM Trans. Graph.*, 21(4):807–832, 2002.
- [71] Cha Zhang and Tsuhan Chen. Efficient feature extraction for 2d/3d objects in mesh representation. In *ICIP (3)*, pages 935–938, 2001.
- [72] M. Sabry Hassouna and Aly A. Farag. Variational curve skeletons using gradient vector flow. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(12):2257–2274, 2009.
- [73] Stéphane Lafon, Yosi Keller, and Ronald R. Coifman. Data fusion and multicue data matching by diffusion maps. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1784–1797, 2006.
- [74] Jian Sun, Maks Ovsjanikov, and Leonidas J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Comput. Graph. Forum*, 28(5):1383–1392, 2009.
- [75] Michael M. Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *CVPR*, pages 1704–1711, 2010.

- [76] Yi Fang, Mengtian Sun, Minhyong Kim, and Karthik Ramani. Heat-mapping: A robust approach toward perceptually consistent mesh segmentation. In *CVPR*, pages 2145–2152, 2011.
- [77] Ivan Sipiran and Benjamin Bustos. Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visual Computer*, 27(11):963–976, 2011.
- [78] Andrei Zaharescu, Edmond Boyer, Kiran Varanasi, and Radu Horaud. Surface feature detection and description with applications to mesh matching. In *CVPR*, pages 373–380, 2009.
- [79] Dan Raviv, Michael M. Bronstein, Alexander M. Bronstein, Ron Kimmel, and Nir A. Sochen. Affine-invariant diffusion geometry for the analysis of deformable 3d shapes. In *CVPR*, pages 2361–2367, 2011.
- [80] J. Keustermans D. Smeets P. Suetens C. Maes, T. Fabry and D. Vandermeulen. Feature detection on 3d face surfaces for pose normalisation and recognition. In *BTAS*, 2010.
- [81] Y. Burago D. Burago and S. Ivanov. *A course in metric geometry, Graduate studies in mathematics*. AMS, vol. 33, 2001.
- [82] Isaak Rubinstein and Lev Rubinstein. *Partial Differential Equations in Classical Mathematical Physics*. Cambridge University Press, 1998.
- [83] Schrder P. Wardetzky M. Desbrun M., Grinspun E. Discrete differential geometry: An applied introduction. In *ACM SIGGRAPH ASIA Courses.*, 13(2), 2008.
- [84] Hao Zhang, Oliver van Kaick, and Ramsay Dyer. Spectral mesh processing. *Comput. Graph. Forum*, 29(6):1865–1894, 2010.

- [85] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace-spectra as fingerprints for shape matching. In *SPM '05: Proceedings of the 2005 ACM Symposium on Solid and Physical Modeling*, pages 101–106, 2005.
- [86] Max Wardetzky, Saurabh Mathur, Felix Kälberer, and Eitan Grinspun. Discrete laplace operators: no free lunch. In *Symposium on Geometry Processing*, pages 33–37, 2007.
- [87] Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *SIGGRAPH*, pages 317–324, 1999.
- [88] Guoliang Xu. Discrete laplace-beltrami operators and their convergence. *Computer Aided Geometric Design*, 21:767–784, 2004.
- [89] Guoliang Xu. Convergent discrete laplace-beltrami operators over triangular surfaces. In *GMP*, pages 195–204, 2004.
- [90] Eitan Grinspun, Yotam Gingold, Jason Reisman, and Denis Zorin. Computing discrete shape operators on general meshes [Eurographics 2006 Best Paper, 3rd Place]. *Eurographics (Computer Graphics Forum)*, 25(3):547–556, 2006.
- [91] Andrew Nealen, Takeo Igarashi, Olga Sorkine, and Marc Alexa. Laplacian mesh optimization. In *GRAPHITE*, pages 381–389, 2006.
- [92] Peter Schröder Alan H. Barr edited by H. C. Hege K. Polthier Mark Meyer, Mathieu Desbrun. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and Mathematics III*, 27(2):113–134, 2002.
- [93] Artiom Kovnatsky, Michael M. Bronstein, Alexander M. Bronstein, Klaus Glashoff, and Ron Kimmel. Coupled quasi-harmonic bases. *Comput. Graph. Forum*, 32(2):439–448, 2013.

- [94] Olgierd Cecil Zienkiewicz and Robert Leroy Taylor. *The finite element method*, volume 3. McGraw-hill London, 1977.
- [95] Elton P. Hsu. *Stochastic Analysis on Manifolds*. American Mathematical Society, 2002.
- [96] Meng Yang, Lei Zhang, David Zhang, and Shenlong Wang. Relaxed collaborative representation for pattern classification. In *CVPR*, pages 2224–2231, 2012.
- [97] Lei Zhang, Meng Yang, and Xiangchu Feng. Sparse representation or collaborative representation: Which helps face recognition? In *ICCV*, pages 471–478, 2011.
- [98] Zhouhui Lian, Afzal Godil, Benjamin Bustos, Mohamed Daoudi, Jeroen Hermans, Shun Kawamura, Yukinori Kurita, Guillaume Lavoué, Hien Van Nguyen, Ryutarou Ohbuchi, Yuki Ohkita, Yuya Ohishi, Fatih Porikli, Martin Reuter, Ivan Sipiran, Dirk Smeets, Paul Suetens, Hedi Tabia, and Dirk Vandermeulen. Shrec '11 track: Shape retrieval on non-rigid 3d watertight meshes. In *3DOR*, pages 79–88, 2011.
- [99] Ke Huang and Selin Aviyente. Sparse representation for signal classification. In *NIPS*, pages 609–616, 2006.
- [100] Philip Shilane, Patrick Min, Michael M. Kazhdan, and Thomas A. Funkhouser. The princeton shape benchmark. In *SMI*, pages 167–178, 2004.
- [101] P.E. Hart R.O. Duda and D.G. Stork. *Pattern Classification*. New York: John Wiley & Sons,, 2001.
- [102] Katarzyna Gebal, Jakob Andreas Bærentzen, Henrik Aanaes, and Rasmus

- Larsen. Shape analysis using the auto diffusion function. *Comput. Graph. Forum*, 28(5):1405–1413, 2009.
- [103] Yong-Jin Liu, Yi-Fu Zheng, Lu Lv, Yuming Xuan, and Xiaolan Fu. 3d model retrieval based on color + geometry signatures. *The Visual Computer*, 28(1):75–86, 2012.
- [104] Andrei Zaharescu, Edmond Boyer, and Radu Horaud. Keypoints and local descriptors of scalar functions on 2d manifolds. *International Journal of Computer Vision*, 100(1):78–98, 2012.
- [105] Jose A. Iglesias and Ron Kimmel. Schrödinger diffusion for shape analysis with texture. In *ECCV Workshops (1)*, pages 123–132, 2012.
- [106] Silvia Biasotti, Andrea Cerri, Daniela Giorgi, and Michela Spagnuolo. Phog: Photometric and geometric functions for textured shape retrieval. *Comput. Graph. Forum*, 32(5):13–22, 2013.
- [107] Alexander Grigoryan.
- [108] Silvia Biasotti, Andrea Cerri, Daniela Giorgi, and Michela Spagnuolo. Phog: Photometric and geometric functions for textured shape retrieval. *Comput. Graph. Forum*, 32(5):13–22, 2013.
- [109] W. Zhao, R. Chellappa, and A. Rosenfeld. Face recognition: a literature survey. *ACM Computing Surveys*, 35:399–458, 2003.
- [110] Andrea F. Abate, Michele Nappi, Daniel Riccio, and Gabriele Sabatino. 2d and 3d face recognition: A survey. *Pattern Recogn. Lett.*, 28:1885–1906, 2007.
- [111] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, December 2003.

- [112] P.J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002. Technical report, <http://www.frvt.org>, March 2003.
- [113] J. Kittler, A. Hilton, M. Hamouz, and J. Illingworth. 3D assisted face recognition: A survey of 3D imaging, modelling and recognition approaches. In *Proc. of CVPR*, 2005.
- [114] Alize E. H. Scheenstra, Arnout Ruifrok, and Remco C. Veltkamp. A survey of 3d face recognition methods. In *Proce. of Audio- and Video-Based Biometric Person Authentication*, volume 3546, pages 891–899, 2005.
- [115] Kevin W. Bowyer, Kyong Chang, and Patrick Flynn. A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition. *Comput. Vis. Image Underst.*, 101(1):1–15, 2006.
- [116] C. Heshner, A. Srivastava, and G. Erlebacher. A novel technique for face recognition using range imaging. In *Proc. of 7th Int. Symp. on Sig. Proc. and Its App.*, pages 201–204, 2003.
- [117] A. Ahmed, A. Farag, and T. Starr. A new statistical model combining shape and spherical harmonics illumination for face reconstruction. In *Proc. of 3rd ISVC-07*, pages 531–541, 2007.
- [118] U. Park and A. K. Jain. 3D face reconstruction from stereo video. In *Proc. of First International Workshop on Video Processing for Security*, 2006.
- [119] Adrian Broadhurst, Tom Drummond, and Roberto Cipolla. A probabilistic framework for space carving. In *Proc. of ICCV*, pages 388–393, 2001.
- [120] S. Hassan Amin and Duncan Gillies. Analysis of 3D face reconstruction. In *Proc. of International Conference on Image Analysis and Processing*, pages 413–418, 2007.

- [121] Volker Blanz and Thomas Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Trans. on PAMI*, 25(9):1063–1074, 2003.
- [122] Dirk Smeets, Johannes Keustermans, Jeroen Hermans, Peter Claes, Dirk Vandermeulen, and Paul Suetens. Symmetric surface-feature based 3d face recognition for partial data. In *IJCB*, pages 1–6, 2011.
- [123] Ioannis A. Kakadiaris, Georgios Passalis, George Toderici, Mohammed N. Murtuza, Yunliang Lu, Nikolaos Karampatziakis, and Theoharis Theoharis. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(4):640–649, 2007.
- [124] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30, 2005.
- [125] David Cristinacce and Timothy F. Cootes. Facial feature detection using adaboost with shape constraints. In *BMVC*, pages 1–10, 2003.
- [126] Mark Everingham, Josef Sivic, and Andrew Zisserman. Hello! my name is... buffy” – automatic naming of characters in tv video. In *BMVC*, pages 899–908, 2006.
- [127] Paul A. Viola and Michael J. Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR (1)*, pages 511–518, 2001.
- [128] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. *ECCV*, 2008.
- [129] Shalini Gupta, K. R. Castleman, Mia K. Markey, and Alan C. Bovik. Texas 3d face recognition database. <http://live.ece.utexas.edu/research/texas3dfr/index.htm>.

- [130] Kyong I. Chang, Kevin W. Bowyer, and Patrick J. Flynn. An evaluation of multimodal 2d+3d face biometrics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):619–624, 2005.
- [131] Chiraz BenAbdelkader and Paul A. Griffin. Comparing and combining depth and texture cues for face recognition. *Image Vision Comput.*, 23(3):339–352, 2005.
- [132] Xiaoguang Lu, Anil K. Jain, and Dirk Colbry. Matching 2.5d face scans to 3d models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(1):31–43, 2006.
- [133] Hossam Abdelmunim, Dongqing Chen, Aly A. Farag, Ross Pusateri, Cambron N. Carter, Mike Miller, Allan G. Farman, and David Tasman. A 3d human teeth database construction based on a point-based shape registration. In *ICIP*, pages 1617–1620, 2011.



## Appendix A: The Derivation of the Heat Equation

In this appendix, seeking to derive the heat equation (1D case) from the physical properties a solution for the rate of heat flow through an object will be given, first from the properties of the object, and second by measuring the rate of heat flow through the boundaries of the object.

Experimental calculations show that the heat  $Q$  in a small volume  $\Delta V$  at time  $t$  can be defined by:

$$\Delta Q = c\rho U \Delta V \quad (\text{A-1})$$

where  $c$  is the specific heat,  $\rho$  is the density, and  $U$  is the temperature. Consider a thin bar of a homogenous material and perfectly insulated along its length so that heat can only flow through its ends. Any position along the bar is denoted as  $x$ , and the length of the bar is denoted as  $L$  such that  $0 \leq x \leq L$ . The temperature  $U$  depends on the position  $x$  and time  $t$ . Thus

$$\Delta Q = c\rho U(x,t) \Delta V \quad (\text{A-2})$$

Now consider a small section of the bar  $U$  defined as the interval from

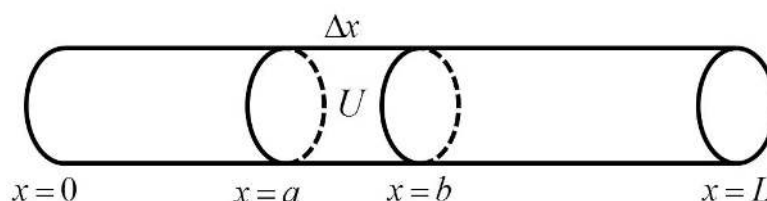


Figure A-1: Thin bar of a homogenous material, any position along the bar is denoted as  $x$ , and the length of the bar is denoted as  $L$  such that  $0 \leq x \leq L$ .

$x = a$  to  $x = b$  as shown in Figure A-1.

The cross sectional area is defined as  $S$ , and the width of this section is  $\Delta x$ . This gives  $\Delta V = S\Delta x$ .

The amount of heat in the cross-sectional area can be expressed as:

$$\Delta Q = c\rho U(x, t)S\Delta x \quad (\text{A-3})$$

Taking the integral to find the amount of heat in the section  $U$  at time  $t$ , it will be:

$$Q(t) = \int_a^b c\rho U(x, t)Sdx \quad (\text{A-4})$$

Since the rod has uniform thickness,  $S$  doesn't change with respect to time, and assume dealing with homogenous materials  $c$  and  $\rho$  do not change with respect to time. Thus, by differentiating and taking the partial of  $U$  to find the change in heat with respect to time, it will be:

$$\frac{dQ}{dt} = \int_a^b c\rho \frac{\partial U}{\partial t} Sdx \quad (\text{A-5})$$

Now from the rate of heat flow through the boundaries of the object another equation for  $\frac{dQ}{dt}$  can be driven.

Since the rate of heat flow through the bar is inversely proportional to the width  $\Delta x$  of  $U$ , and directly proportional to the cross-sectional area.

$$\Delta Q = -\mathcal{C} \frac{U(a + \Delta x, t) - U(a, t)}{\Delta x} S \quad (\text{A-6})$$

The proportionality constant  $\mathcal{C}$  is known as the thermal conductivity. Letting  $\Delta x \rightarrow 0$  in Equation A-6, then the rate of heat flow through  $U$  at  $x = a$  is given by

$$-\mathcal{C} \frac{\partial U}{\partial x}(a, t) S \quad (\text{A-7})$$

Following the same argument it can be shown that the rate of heat flow through  $U$  at  $b$  is defined as:

$$-\mathcal{C} \frac{\partial U}{\partial x}(b, t) S \quad (\text{A-8})$$

Therefore, the amount of heat that  $U$  obtains at time  $t$  can be given by

$$\frac{dQ}{dt} = \mathcal{C} \left[ \frac{\partial U}{\partial x}(b, t) - \frac{\partial U}{\partial x}(a, t) \right] S \quad (\text{A-9})$$

Or,

$$\frac{dQ}{dt} = \int_a^b \frac{d}{dx} \left( \mathcal{C} \frac{\partial U}{\partial x} S \right) dx \quad (\text{A-10})$$

Since the material of the bar is homogeneous and the cross-sectional area is constant, so  $\mathcal{C}$  and  $S$  to be constant. Therefore the expression becomes

$$\frac{dQ}{dt} = \mathcal{C} \int_a^b \frac{\partial^2 U}{\partial x^2} S dx \quad (\text{A-11})$$

From Equation A-5, and Equation A-11

$$c\rho \int_a^b \frac{\partial U}{\partial t} S dx = \mathcal{C} \int_a^b \frac{\partial^2 U}{\partial x^2} S dx \quad (\text{A-12})$$

$$c\rho \int_a^b \frac{\partial U}{\partial t} dx - \mathcal{C} \int_a^b \frac{\partial^2 U}{\partial x^2} dx = 0, \quad (\text{A-13})$$

$$\frac{\partial U}{\partial t} - k \frac{\partial^2 U}{\partial x^2} = 0, \quad (\text{A-14})$$

where  $k = \mathcal{C}/c\rho$  and it is called thermal diffusivity. Now the heat equation

is developed, also known as the diffusion equation. This equation models the flow of heat at time  $t$ . The heat equation is a second order parabolic partial differential equations, and usually written as

$$U_t = kU_{xx}, \tag{A-15}$$

□

## Appendix B: Analytic Derivation of Equation 3.32

To discretize (3.29) let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Choose linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  is equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolates between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $t_k$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at vertices  $i, j, k$  respectively, and  $|t_k|$  is the area of  $t_k$ , as shown in Figure B-1, then it can be shown that:

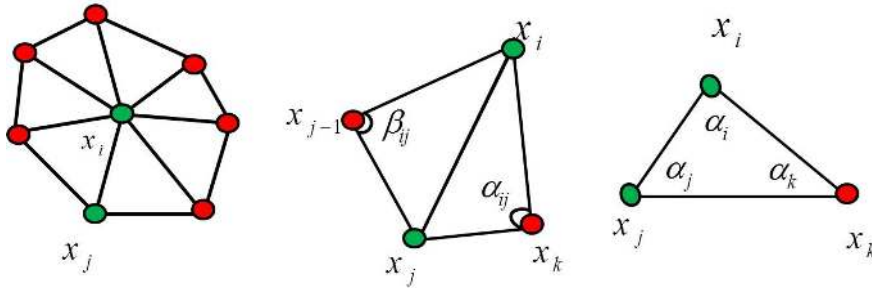


Figure B-1: Discrete triangulated surface. Left: A vertex  $\mathbf{x}_i$  and its adjacent faces Middle: the definition of the angles  $\alpha_{ij}$  and  $\beta_{ij}$  Right: the definition of the interior angles  $\alpha_i, \alpha_j, \alpha_k$  of triangle  $T$  at vertices  $i, j, k$  respectively.

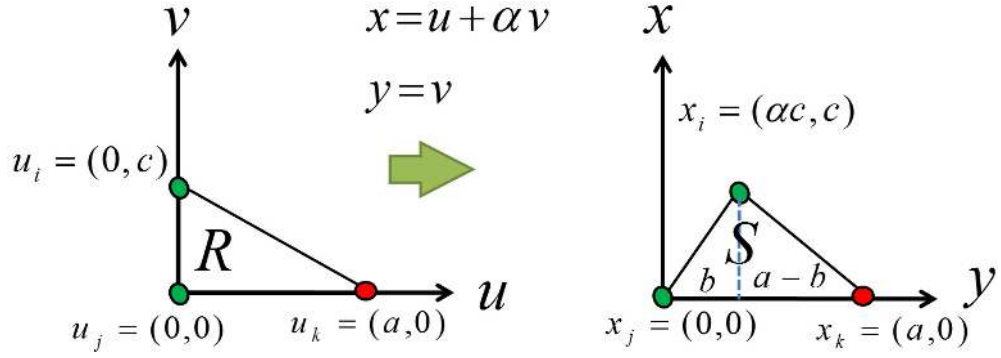


Figure B-2: Shearing a regular triangle to a right angle triangle by changing variables

$$\nabla\varphi_i \cdot \nabla\varphi_j = -\frac{\cot \alpha_k}{2 |t_k|} \quad (\text{B-1})$$

$$|\nabla\varphi_i|^2 = \frac{\cot \alpha_j + \cot \alpha_k}{2 |t_k|} \quad (\text{B-2})$$

$$\int_T \varphi_i d\sigma = \frac{|t_k|}{3} \quad (\text{B-3})$$

$$\int_T \varphi_i^2 d\sigma = \frac{|t_k|}{6} \quad (\text{B-4})$$

$$\int_T \varphi_i \varphi_j d\sigma = \frac{|t_k|}{12} \quad (\text{B-5})$$

$$(\text{B-6})$$

For simplicity, assume a shearing transformation is applied on a general triangle  $S$  to get a right angle triangle  $R$  as shown in Figure B-2, then define  $\varphi_i := \frac{y}{c}$  so it will be 1 at vertex  $i$  and 0 on all triangles incident, and  $\varphi_j := \frac{-b}{ac}(y - \frac{b}{c}x)$  so it will be 1 at vertex  $j$  and 0 on all triangles incident, then  $\nabla\varphi_i = (0, \frac{1}{c})$ , and  $\nabla\varphi_j = (\frac{1}{a}, -\frac{b}{ac})$ , and the product  $\langle \nabla\varphi_i \cdot \nabla\varphi_j \rangle$  will be:

$$\langle \nabla\varphi_i \cdot \nabla\varphi_j \rangle = \frac{-b}{ac^2} = \frac{-\frac{b}{c}}{ac} = \frac{-\cot \alpha_k}{ac} = -\frac{\cot \alpha_k}{2 |t_k|} \quad (\text{B-7})$$

this proves the first part.

$$\text{Also, } |\nabla\varphi|^2 = \frac{1}{c^2} = \frac{b+a-b}{ac^2} = \frac{\frac{b}{c} + \frac{a-b}{c}}{ac}$$

this implies that:

$$|\nabla\varphi|^2 = \frac{\cot\alpha_j + \cot\alpha_k}{2|t_k|}$$

where  $|t_k|$  is the area of  $R$  this proves the second part.

To do the proof of the other parts of 3.32 the integration by substitution will be used, also known as u-substitution, is a method for finding integrals. Using the fundamental theorem of calculus often requires finding an antiderivative.

Let  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a function defined on a triangle  $S$ , a shearing transformation will be applied on  $S$  to get a right angle triangle  $R$  as shown in Figure B-2, then:

$$\int_S f(x, y) dx dy = \int_R f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv \quad (\text{B-8})$$

if  $\alpha c = b$  then  $x = u + \alpha v = u + \frac{b}{c}v$ , and  $y = v$

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \begin{vmatrix} 1 & \frac{b}{c} \\ 0 & 1 \end{vmatrix} = 1$$

$$\begin{aligned} \int_S \varphi_i(x, y) dx dy &= \frac{1}{c} \int_S y dx dy \\ &= \frac{1}{c} \int_R v du dv \\ &= \frac{1}{c} \int_0^c \int_0^{a-\frac{a}{c}v} v du dv \\ &= \frac{1}{c} \int_0^c v(a - \frac{a}{c}v) dv \\ &= \frac{1}{c} \left( \frac{ac^2}{2} - \frac{ac^3}{3c} \right) \\ &= \left( \frac{ac}{2} - \frac{ac}{3} \right) \\ &= \frac{ac}{6} \\ &= \frac{|t_k|}{3} \end{aligned}$$

(B-9)

$$\begin{aligned}
\int_S \varphi_i(x, y)^2 dx dy &= \frac{1}{c^2} \int_S y^2 dx dy \\
&= \frac{1}{c^2} \int_R v^2 du dv \\
&= \frac{1}{c^2} \int_0^c \int_0^{a-\frac{a}{c}v} v^2 du dv \\
&= \frac{1}{c^2} \int_0^c v^2 (a - \frac{a}{c}v) dv \\
&= \frac{1}{c^2} \left( \frac{ac^3}{3} - \frac{ac^4}{4c} \right) \\
&= \left( \frac{ac}{3} - \frac{ac}{4} \right) \\
&= \frac{ac}{12} \\
&= \frac{|t_k|}{6}
\end{aligned}$$

(B-10)

$$\begin{aligned}
\int_S \varphi_i(x, y) \varphi_j(x, y) dx dy &= \frac{-b}{ac^2} \int_S y(y - \frac{b}{c}x) dx dy \\
&= \frac{-b}{ac^2} \int_R v(v - \frac{c}{b}(u + \frac{b}{c}v)) \cdot 1 \cdot du dv \\
&= \frac{-b}{ac^2} \int_R v(v - \frac{c}{b}u - v) \cdot 1 \cdot du dv \\
&= \frac{1}{ac} \int_R uv du dv \\
&= \frac{1}{ac} \int_0^c \int_0^{a-\frac{a}{c}v} uv du dv \\
&= \frac{1}{2ac} \int_0^c v(a - \frac{a}{c}v)^2 dv \\
&= \frac{1}{2ac} \left( \frac{a^2 c^2}{2} - 2 \frac{a^2 c^3}{3c} + \frac{a^2 c^4}{4c^2} \right) \\
&= \frac{ac}{24} \\
&= \frac{|t_k|}{12}
\end{aligned}$$

(B-11)



## Appendix C: Analytic Derivation of Equations 3.34 and 3.34

Let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Choose linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  are equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolates between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $t_k$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at vertices  $i, j, k$  respectively, and  $|t_k|$  is the area of  $t_k$ , as shown in Figure B-1. If  $i$  and  $j$  are distinct nonadjacent vertices, then  $B(i, j) = L(i, j) = 0$ . If  $i$  and  $j$  are adjacent vertices then:

$$B(i, j) = \int_{\mathcal{M}} \varphi_i \varphi_j d\sigma,$$

Assuming that integration over  $\mathcal{M}$  is equal to summing the integrations over all triangles, substituting from Equation B-1, then:

$$\begin{aligned} B(i, j) &= \sum_{i, j} \int_{t_k} \varphi_i \varphi_j d\sigma \\ &= \sum_{i, j} \frac{|t_k|}{12} \end{aligned}$$

and,

$$\begin{aligned} B(i, i) &= \sum_{i,i} \int_{t_k} \varphi_i \varphi_i d\sigma \\ &= \sum_{i,i} \frac{|t_k|}{6} \end{aligned}$$

so

$$B(i, j) = \begin{cases} \frac{|t_r|+|t_s|}{12} & \text{if } j \in N_e(i) \\ \frac{\sum_{k \in N_e(i)} |t_k|}{6} & \text{if } i = j \end{cases} \quad (\text{C-1})$$

Similarly,

$$L(i, j) = \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) d\sigma.$$

then:

$$\begin{aligned} L(i, j) &= \sum_{i,j} \int_{t_k} \nabla \varphi_i \cdot \nabla \varphi_j d\sigma \\ &= \sum_{i,j} |t_k| \left( -\frac{\cot \alpha_k}{2 |t_k|} \right) \end{aligned}$$

by substitution from Equation B-1, then:

$$L(i, j) = \begin{cases} w(i, j) = \frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2} & \text{if } j \in N_e(i) \\ -\sum_{k \in N_e(i)} w(i, k) & \text{if } i = j \\ 0 & \text{else} \end{cases} \quad (\text{C-2})$$

□

## Appendix D: Analytic Derivation of Equation 5.10

This appendix derives the analytic expression of the 5.10

To discretize (5.9) let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Choose linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  are equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolates between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $T$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at vertices  $i, j, k$  respectively, and  $A(T)$  is the area of  $T$ , as shown in Figure B-1,

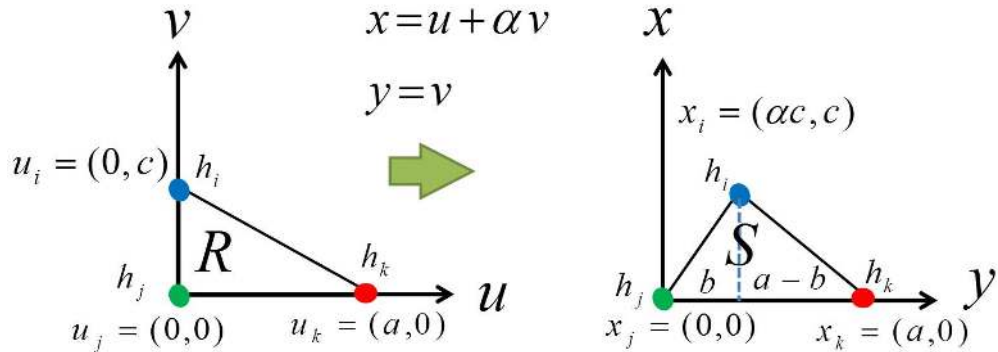


Figure D-1: Shearing a regular triangle to a right angle triangle by changing variables.

then it can be shown that:

$$\begin{aligned}
\nabla\varphi_i \cdot \nabla\varphi_j &= -\frac{\cot\alpha_k}{2A(T)} \\
|\nabla\varphi|^2 &= \frac{\cot\alpha_j + \cot\alpha_k}{2A(T)} \\
\int_T \varphi_i d\sigma &= \frac{A(T)}{3} \\
\int_T \varphi_i^2 d\sigma &= \frac{A(T)}{6} \\
\int_T \varphi_i \varphi_j d\sigma &= \frac{A(T)}{12} \\
\int_T \varphi_i^3 d\sigma &= \frac{A(T)}{10} \\
\int_T \varphi_i^2 \varphi_j d\sigma &= \frac{A(T)}{30} \\
\int_T \varphi_i \varphi_j \varphi_k d\sigma &= \frac{A(T)}{60}
\end{aligned}
\tag{D-1}$$

The first four parts of 5.10 are given in Appendix B. Similarly and referring to Figure D-1, let  $\varphi_i := \frac{y}{c}$  so it will be 1 at vertex  $i$  and 0 on all triangles incident,  $\varphi_j := \frac{-b}{ac}(y - \frac{b}{c}x)$  so it will be 1 at vertex  $j$  and 0 on all triangles incident, and  $\varphi_k := \frac{b-a}{ac}(y - \frac{b}{b-a}(x - a))$  so it will be 1 at vertex  $j$  and 0 on all triangles incident.

$$\begin{aligned}
\int_S \varphi_i(x, y)^3 dx dy &= \frac{1}{c^3} \int_S y^3 dx dy \\
&= \frac{1}{c^3} \int_R v^3 du dv \\
&= \frac{1}{c^3} \int_0^c \int_0^{a-\frac{a}{c}v} v^3 du dv \\
&= \frac{1}{c^3} \int_0^c v^3 (a - \frac{a}{c}v) dv \\
&= \frac{1}{c^3} (\frac{ac^4}{4} - \frac{ac^5}{5c}) \\
&= (\frac{ac}{4} - \frac{ac}{5}) \\
&= \frac{ac}{20} \\
&= \frac{A(T)}{10}
\end{aligned}$$

(D-2)

$$\begin{aligned}
\int_S \varphi_i(x, y)^2 \varphi_j(x, y) dx dy &= \frac{-b}{ac^2} \int_S y^2 (y - \frac{b}{c}x) dx dy \\
&= \frac{-b}{ac^3} \int_R v^2 (v - \frac{c}{b}(u + \frac{b}{c}v)) \cdot 1 \cdot du dv \\
&= \frac{-b}{ac^3} \int_R v^2 (v - \frac{c}{b}u - v) \cdot 1 \cdot du dv \\
&= \frac{1}{ac^2} \int_R uv^2 du dv \\
&= \frac{1}{ac^2} \int_0^c \int_0^{a-\frac{a}{c}v} uv^2 du dv \\
&= \frac{1}{2ac^2} \int_0^c v^2 (a - \frac{a}{c}v)^2 dv \\
&= \frac{1}{2ac^2} (\frac{a^2c^3}{3} - 2\frac{a^2c^4}{4c} + \frac{a^2c^5}{5c^2}) \\
&= \frac{ac}{60} \\
&= \frac{A(T)}{30}
\end{aligned}$$

(D-3)

$$\begin{aligned}
\int_S \varphi_i(x, y) \varphi_j(x, y) \varphi_k dx dy &= \frac{-b(b-a)}{a^2 c^3} \int_S y \left(y - \frac{b}{c} x\right) \left(y - \frac{b}{b-a} (x-a)\right) dx dy \\
&= \frac{1}{ac} \int_R uv du dv - \frac{1}{a^2 c} \int_R uv^2 du dv - \frac{1}{ac^2} \int_R uv^2 du dv \\
&= \frac{ac}{24} - \frac{ac}{60} - \frac{ac}{60} \\
&= \frac{ac}{120} \\
&= \frac{A(T)}{60}
\end{aligned}$$

(D-4)

□

## Appendix E: Analytic Derivation of Equations 5.11, 5.12, and 5.13

To discretize (5.9) let  $\mathcal{N} = (M; \mathcal{T})$  be a triangulated surface that approximates  $\mathcal{M}$ . Here  $M := \{\mathbf{x}_i; i = 1, \dots, n\}$  is a set of  $n$  vertices and  $\mathcal{T}$  is an abstract simplicial complex containing the adjacency information. Choose linearly independent basis functions  $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ , where  $\varphi_i(\mathbf{x}_j) = \delta_{ij}$  are equal to 1 at vertex  $i$ , 0 at all other vertices, and linearly interpolates between 1 and 0 on all triangles incident to vertex  $i$ . Label vertex  $\mathbf{x}_i$  simply as  $i$ . If  $i, j, k$  are the distinct vertices of a triangle  $T$ ,  $\alpha_i, \alpha_j, \alpha_k$  denote the interior angles of  $T$  at vertices  $i, j, k$  respectively, and  $A(T)$  is the area of  $T$ , as shown in Figure B-1. If  $i$  and  $j$  are distinct nonadjacent vertices, then  $B(i, j) = L(i, j) = 0$ . If  $i$  and  $j$  are adjacent vertices then:

$$B(i, j) = \int_{\mathcal{M}} \varphi_i \varphi_j h d\sigma,$$

Assuming that integration over  $\mathcal{M}$  is equal to summing the integrations over all triangles, then:

$$B(i, j) = \sum_{i,j} \int_T \varphi_i \varphi_j h d\sigma$$

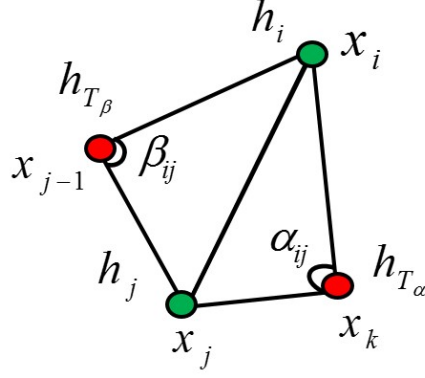


Figure E-1: The definition of the angles  $\alpha_{ij}$  and  $\beta_{ij}$  and the weights  $h_{T_\alpha}$  and  $h_{T_\beta}$ .

and substituting  $h$  by  $h_i\varphi_i + h_j\varphi_j + h_k\varphi_k$ , then:

$$\begin{aligned}
 B(i, i) &= \sum_{i,i} \int_T \varphi_i \varphi_i h d\sigma \\
 &= \sum_{i,i} \int_T \varphi_i \varphi_i (h_i \varphi_i + h_j \varphi_j + h_k \varphi_k) d\sigma \\
 &= \sum_{i,i} \int_T (\varphi_i^2 \varphi_j h_i + \varphi_i \varphi_j^2 h_j + \varphi_i \varphi_j \varphi_k h_k) d\sigma
 \end{aligned}$$

Denote by  $h_T$  the average value of a function  $h$  over a triangle  $T$ . Denote by  $N_e(i)$  the set of vertices adjacent to  $i$ . Given  $j \in N_e(i)$  denote by  $T_\alpha$  and  $T_\beta$  the triangles having  $(i, j)$  as an edge and by  $\alpha$  and  $\beta$  the interior angles of  $T_\alpha$  and  $T_\beta$  opposite edge  $(i, j)$ , as shown in Figure E-1, and substituting from Equation D-1, then:

$$B(i, j) = (h_i + h_j) \frac{A(T_\alpha) + A(T_\beta)}{60} + \frac{h_{T_\alpha} A(T_\alpha) + h_{T_\beta} A(T_\beta)}{20} \quad (\text{E-1})$$

whenever  $j \in N_e(i)$ ,



$$B(i, i) = \sum_{k \in N_e(i)} A(T_k) \left( \frac{h_i}{15} + \frac{h_{T_k}}{10} \right) \quad (\text{E-2})$$

Similarly,

$$L(i, j) = \int_{\mathcal{M}} (\nabla \varphi_i \cdot \nabla \varphi_j) h d\sigma.$$

Substituting  $h$  by  $h_i \varphi_i + h_j \varphi_j + h_k \varphi_k$ , then:

$$\begin{aligned} L(i, j) &= \sum_{i, j} \int_T \nabla \varphi_i \cdot \nabla \varphi_j (h_i \varphi_i + h_j \varphi_j + h_k \varphi_k) d\sigma \\ &= \sum_{i, j} \left( -\frac{\cot \alpha_k}{2A(T)} \right) \int_T (h_i \varphi_i + h_j \varphi_j + h_k \varphi_k) d\sigma \end{aligned}$$

Substituting from Equation D-1, then:

$$L(i, j) = \begin{cases} -\frac{h_{T_\alpha} \cot \alpha + h_{T_\beta} \cot \beta}{2} & \text{if } j \in N_e(i) \\ \sum_{k \in N_e(i)} L(i, k) & \text{if } i = j \end{cases} \quad (\text{E-3})$$

where  $h_{T_\alpha}$ , and  $h_{T_\beta}$  are the average values of a function  $h$  over a triangles  $T_\alpha$ , and  $T_\beta$ .

□

## Appendix F: Iterative Closest Point (ICP)

### Algorithm

The objective of this section is to formulate the registration between two different 3D surface models [133]. Given a source surface model  $\mathbf{M}_s$  and a target shape  $\mathbf{M}_t$ , a transformation  $\mathbf{A}$  that moves points from  $\mathbf{M}_s$  to  $\mathbf{M}_t$  is needed. The selected transformation has a set of parameters will be estimated to minimize a certain energy function. Assume that the transformation is affine and hence it will have the following homogeneous format:

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_6 & a_7 & a_8 \\ a_9 & a_{10} & a_{11} & a_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{F-1})$$

Assume that  $x_i \in \mathbf{M}_s, i = 1, 2, \dots, N_s$  and  $y_i \in \mathbf{M}_t$  is the closest point to  $\mathbf{A}x_i$  on the target shape where  $N_s$  is the number of points on the source surface. Note

that both  $x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ 1 \end{pmatrix}$  and  $y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ 1 \end{pmatrix}$  points are put in the homogeneous

vector notation of size  $4 \times 1$ . Consider the Euclidean distance between the moved point and its closes position to be the dissimilarity measure as follows:

$$d_i = \|\mathbf{A}x_i - y_i\|. \quad (\text{F-2})$$

And hence, the summation of squared differences can be written as the

objective energy function:

$$E = E(a_1, a_2, \dots, a_{12}) = \sum_{i=1}^{N_s} \frac{1}{2} (\mathbf{A}x_i - y_i)^T (\mathbf{A}x_i - y_i) \quad (\text{F-3})$$

Dividing by two just to remove numbers from the derivative equations as follows. The transformation parameters  $\{a_k\}$  are required to minimize the above functional. Taking the derivative of the energy with respect to  $a_k$  will result in:

$$\frac{\partial E}{\partial a_k} = \sum_{i=1}^{N_s} (\mathbf{A}x_i - y_i)^T (\mathbf{A}_{a_k} x_i) \quad (\text{F-4})$$

where  $\mathbf{A}_{a_k}$  is the derivative of the transformation matrix  $\mathbf{A}$  with respect to the parameter  $a_k$ . So, any element  $\mathbf{A}_{a_k}(m, n)$  in row  $m$  and column  $n$  of the derivative matrix can be written in the following format:

$$\mathbf{A}_{a_k}(m, n) = \begin{cases} 1 & \text{if } k == 4(m-1) + n \\ 0 & \text{otherwise} \end{cases} \quad (\text{F-5})$$

Taking derivatives with respect to all the 12 parameters will result in a linear system of equations as shown below:

$$\Psi \Theta = \Lambda \quad (\text{F-6})$$

where  $\Theta = (a_1 a_2 \dots a_{11} a_{12})^T$ . The column vector  $\Lambda$  has 12 elements defined as follows:

$$\Lambda = \sum_{i=1}^{N_s} \begin{pmatrix} x_{i1}y_{i1} \\ x_{i2}y_{i1} \\ x_{i3}y_{i1} \\ y_{i1} \\ x_{i1}y_{i2} \\ x_{i2}y_{i2} \\ x_{i3}y_{i2} \\ y_{i2} \\ x_{i1}y_{i3} \\ x_{i2}y_{i3} \\ x_{i3}y_{i3} \\ y_{i3} \end{pmatrix} \quad (\text{F-7})$$

The square matrix  $\Psi$  has  $12 \times 12$  elements in Eq. F-8.

$$\Psi = \sum_{i=1}^{N_s} \begin{pmatrix} X_{ij} & 0 & 0 \\ 0 & X_{ij} & 0 \\ 0 & 0 & X_{ij} \end{pmatrix} \quad (\text{F-8})$$

Where

$$X_{ij} = \begin{pmatrix} x_{i1}^2 & x_{i1}x_{i2} & x_{i1}x_{i3} & x_{i1} \\ x_{i1}x_{i2} & x_{i2}^2 & x_{i2}x_{i3} & x_{i2} \\ x_{i1}x_{i3} & x_{i2}x_{i3} & x_{i3}^2 & x_{i3} \\ x_{i1} & x_{i2} & x_{i3} & 1 \end{pmatrix} \quad (\text{F-9})$$

Solving the above linear system will compute the parameters of the affine transformation that minimizes the distance between the two shape models. This process is repeated on the transformed surface until the change in the transformed points is not significant.

# NOMENCLATURE

## List of Abbreviations

BoF	Bag of Features
BRIEF	Binary Robust Independent Elementary Features
BRISK	Binary Robust Invariant Scalable Keypoints
CAD	Computer Aided Design
CRC-RLS	Collaborative Representation-Based Classification with Regularized Least Square
DCG	Discounted Cumulative Gain
DOG	Difference of Gaussians
FP	False Positive rate
HKS	Heat Kernel Signatures
HOG	Histogram of Gradients
LBP	Local Binary Patterns
LIOP	Local Intensity Order Pattern for Feature Description
NN	Nearest Neighbor
ORB	Oriented FAST and Rotated BRIEF
SI-HKS	Scale Invariant Heat Kernel Signatures

SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
TP	True Positive rate
W-HKS	Weighted Heat Kernel Signatures

### List of Symbols

$E, F, G$	the three first fundamental forms
$(\mathbf{u}, \mathbf{v})$	an orthogonal basis on a tangent plane of $\mathbf{S}$
$\alpha_i, \alpha_j, \alpha_k$	Interior angles of $T$ at vertices $i, j, k$ respectively
$\Delta_{\mathcal{M}}f$	Laplace-Beltrami operator
$\frac{dQ}{dt}$	The change in heat with respect to time
$\gamma$	the path of an interval
$\kappa_1$	maximum principal curvature
$\kappa_2$	minimum principal curvature
$\langle \cdot, \cdot \rangle$	Inner product
$\mathbb{R}$	The real numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}^2$	2D real number Cartesian space
$\mathbb{R}^3$	3D real number Cartesian space
$\mathbf{S}$	a compact, regular and orientable 3D surface
$\mathbf{X}$	The metric space

$\mathbf{x} = (x, y, z)$	a vertex on the 3D surface $\mathbf{S}$
$\mathcal{B} = \{\varphi_i\}_{i=1}^n$	Linearly independent basis functions
$\mathcal{C}$	Class of r-times continuously differentiable maps
$\mathfrak{c}$	The thermal conductivity
$\mathcal{M}$	Riemannian manifold
$c$	The specific heat
$\mu$	A measure over manifold $\mathcal{M}$
$\nabla f$	Gradient of function $f$
$\rho$	The density
$L, M, N$	the three second fundamental forms
$\Delta Q$	The heat $Q$ in a small volume $\Delta V$ at time $t$
$\Delta V$	Small volume
$A(T)$	The area of triangle $T$
$A^T$	Transpose of matrix $A$
$a_{ij}$	ijth element of matrix $A$
$B(i, j)$	The weight matrix
$d_L$	metric induced by the length
$\det(A)$	Determinant of matrix $A$
$g$	The Riemannian metric
$H$	mean curvature

$K$	Gaussian curvature
$L$	length of the path
$L(i, j)$	The laplacian matrix
$Q$	The heat
$trace(A)$	Trace of matrix A
$U$	The temperature



# CURRICULUM VITAE

MOSTAFA ABDELRAHMA

## Contact Information

Department of Electrical and Computer Engineering  
University of Louisville  
Computer Vision and Image Processing Laboratory  
2211 South Brook Lutz Hall, Room # 414,  
Louisville, KY 40292, USA

*Mobile:* +1-502-296-8962

*E-mail:*

maabde01@louisville.edu

mostafa.a.abdelrahman@gmail.com

www.cvip.uofl.edu

## Objective

Pursuit of advanced research and development in computer vision, image processing, machine learning as a postdoctoral, research scientist, faculty or being hired by a leading edge company conducting research and development that allows for advanced research in computer vision, machine learning.

## Research Interests

**Theoretical and computational methods in metric geometry and their application to problems in computer vision, pattern recognition, shape analysis, computer graphics, image processing, and machine learning.** Also facial recognition under varying pose, illumination, occlusion and expression.

## Education

**University of Louisville**, Louisville, Kentucky, USA

- Ph.D. Candidate,

Electrical and Computer Engineering, Expected to Graduate by Fall 2013.

- Dissertation Topic: *GEOMETRIC MODELING OF NON-RIGID 3D SHAPES: THEORY AND APPLICATION TO OBJECT RECOGNITION*

- GPA : 4 out of 4

**Assiut University**, Assiut, Egypt

- M.Sc.,  
Faculty of Eng., Electrical & Computer Engineering Dep. June , 2006
  - Thesis Topic: *Design of Different Compensators for Double Integrator Plants*
- B.Sc.,  
Faculty of Eng., Electrical & Computer Engineering Dep., May 2001
  - GPA: 4 out of 4 (Distinction With Honors - First of Class)
  - Project: Design of Multi-drop Serial Industrial Bus

## Academic Appointments

**Graduate Research Assistant** January 2009 to present

- CVIP Lab. Department of Electrical and Computer Engineering, University of Louisville
- **Research and Development:**
  - Shape analysis and classification: develop a flexible and efficient methods for shape representation, especially for non-rigid 3D shapes.
  - Face recognition and identification at a distance for indoor and outdoor environments (2D and 3D).
  - Human face 3D reconstruction from stereo Images.
  - Remote Sensing for Surveillance of Critical Infrastructure.
  - Human jaw reconstruction and analysis from optical imaging for orthodontic treatments
  - Distinction between autistic corpus callosum and normal ones for autism diagnosis

**Graduate Research Assistant** May 2003 to December 2008

- Faculty of Eng., Electrical & Computer Engineering Dep.,  
**Assiut University**, Assiut, Egypt
- **Research and Development:**
  - Control System Design using optimal time, optimal energy, adaptive fuzzy logic controller, and genetic optimization.

## Other Working Experience

### Internship

May 2010 - July 2010 (3 months)

- DynaVox Technologies,
- **Activities:**
  - Facial key points extraction

### Part Time Trainer

August 2007 - December 2009

- Siemens Control Systems Laboratory, Assiut University
- **Activities:**
  - Teaching Siemens SIMATIC Manager, PCS7, WINCC, PDM, Sensors and Communication, and Standard, and Drives inverter (MM4).

## Apprenticeships and Training

### Siemens Control Systems

July 2007, Siemens Egypt.

- **Content:**
  - Siemens SIMATIC Manager, PCS7, WINCC, PDM, Sensors and Communication, and Standard, and Drives inverter (MM4).

## Publications

- **M. Abdelrahman**, M. T. El-Melegy, A. Farag: Heat Front Propagation Contours for 3D Face Recognition. Proc. of IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS 2013).
- **M. Abdelrahman**, M. T. El-Melegy, A. Farag: A Novel Approach for 3D Object Classification using Sparse Scale Invariant Heat Kernels. Pattern Recognition Letters, (2013) under review.
- A. Cerri, S. Biasotti, **M. Abdelrahman**, J. Angulo, K. Berger, L. Chevalier, M. El-Melegy, A. Farag, F. Lefebvre, A. Giachetti, H. Guermoud, Y.-J. Liu, S. Velasco-Forero, JR. Vigouroux, C.-X. Xu, J.-B. Zhang, SHREC'13 Track: Retrieval on textured 3D models. Eurographics Workshop on 3D Object Retrieval (2013).
- J. Machado, A. Ferreira, P. B. Pascoal, **M. Abdelrahman**, M. Aono, M. El-Melegy, A. Farag, H. Johan, B. Li, Y. Lu, A. Tatsuma, SHREC 2013: Retrieval of objects captured with low-cost depth-sensing cameras, Eurographics Workshop on 3D Object Retrieval (2013).

- **M. Abdelrahman**, M. El-Melegy, A. Farag: 3D Object Classification using Scale Invariant Heat Kernels with Collaborative Classification. NORDIA ECCV workshop 2012.
- **M. Abdelrahman**, M. El-Melegy, A. Farag: Heat Kernels for Non-Rigid Shape Retrieval: Sparse Representation and Efficient Classification. CRV 2012: 153-160.
- **M. Abdelrahman**, A. Ali, S. Elhabian, H. Rara, and A. Farag. A Passive Stereo System for 3D Human Face Reconstruction and Recognition at a Distance. IEEE CVPR Workshop on Biometrics, pp. 17 - 22, June 2012.
- **M. Abdelrahman**, A. Ali, A. Farag, M. Casanova, A. Farag: New Approach for Classification of Autistic vs. Typically Developing Brain Using White Matter Volumes. CRV 2012: 284-289.
- A. Abdelrehim, **M. Abdelrahman**, H. Abdelmunim, A. Farag, and M. Miller, Novel Image-Based 3D Reconstruction of the Human Jaw using Shape from Shading and Feature Descriptors, British Machine Vision Conference, (BMVC), vol. 41, pp. 1-11, Aug 2011.
- **M. Abdelrahman**, S. Elhabian, A. Ali and A. Farag, Face Recognition at-a-Distance using Texture, Sparse-Stereo, and Dense-Stereo, International Conference on Multimedia Technology (ICMT), pp. 6690-6695, 26-28 July 2011.
- **M. Abdelrahman**, A. Ali, S. Elhabian and A. Farag. Solving Geometric Co-Registration Problem of Multi-Spectral Remote Sensing Imagery Using SIFT-based Features Toward Precise Change Detection. 7th International Symposium on Visual Computing (ISVC). Las Vegas, Nevada, USA, September 26-28, 2011, vol.2, pp 607-616.
- **M. Abdelrahman**, A. Ali, and A. Farag, "Precise change detection in multi-spectral remote sensing imagery using SIFT-based registration," IEEE International Conference on Multimedia Technology (ICMT), pp. 6238 - 6242, July 2011.
- A. Abdelrahim, **M. Abdelrahman**, A. Farag, "Image retrieval based on content and image compression," IEEE International Conference on Multimedia Technology (ICMT), pp. 6696 - 6703 , July 2011.
- A. Farag, S. Elhabian, **M. Abdelrahman**, J. Graham, A. Farag, D. Chen, and M. F. Casanova. Surface Modeling of the Corpus Callosum. Proc. of International Symposium on Visual Computing (ISVC), pp. 9-18, 2010.
- A. Farag, S. Elhabian, **M. Abdelrahman**, J. Graham, A. Farag, D. Chen, and M. F. Casanova. Shape Modeling of the Corpus Callosum. Proc. of the 32nd IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4288-4291, 2010.

- A. I. Saleh ; M. M. Hasan ; A. A. Ali, and **M. Abdelrahman**. "Design Of Different Compensators For Double Integrator Plants ", Journal of Engineering Science, vol. 34, no. 3, PP. 937-958, May 2006.

### **Referee Service**

- Reviewer at IET Computer Vision

### **Honors and Awards**

- Egyptian Engineering Union award 2001.
- 6th October University-Egypt honoring.
- Honorable mention, 5th Annual Graduate Student Research Symposium, March 21, 2013.
- Electrical Engineering Outstanding Graduate Student Award 2013.
- Theobald Scholarship Award 2013.

### **Teaching Experience**

#### **University of Louisville**, Louisville, Kentucky USA

- ECE 620: Pattern Recognition
- ECE 619: Computer Vision
- Materials are available at CVIP Lab website
- ECE 550/551: Communications/Lab
- WKU 470/475: Western Kentucky University Communications/Lab.

#### **Assiut University** Assiut, Egypt

- Helped in teaching the following courses:
  - Computer Interface, Microprocessor, Embedded Systems Design, Digital Signal Processing, Digital Circuits Design, Advanced Mathematics, Fuzzy Logic and Neural Networks, and C++ programming.
- Helped in advising several graduation projects for B.Sc. students.
  - Digital control of DC motor.
  - Design of Home Automation system using PIC and PC.
  - Controlling of Mixing System using PLC, HMI (2006-2007)

## Skills

### Programming Languages and Toolkits

- Compiled: C, C++, C#.
- Scripted: MatLab, Maple, Unix shell scripting.
- Hardware-level: VHDL, Assembly languages.
- Libraries and tools: EMGU (C#), OpenCV (C++), PostgreSQL, Point Cloud Library (PCL), Parallel Programming with CUDA, Visualization Toolkit (VTK)

### Others

- Presentations preparation skills
- Writing skills
- Course preparation skills
- Instructing skills

## Membership

- IEEE student member since 2011.

## Languages

- |           |                |
|-----------|----------------|
| • Arabic  | Mother Tongue  |
| • English | Fluent - TOEFL |
| • French  | Basic          |