Geometric Analysis of Shapes and Its application to Medical Image Analysis

by

Anirban Mukhopadhyay

(Under the DIRECTION of Suchendra M. Bhandarkar)

Abstract

Geometric analysis of shapes plays an important role in the way the visual world is perceived by modern computers. To this end, low-level geometric features provide most obvious and important cues towards understanding the visual scene. A novel intrinsic geometric surface descriptor, termed as the Geodesic Field Estimate (GFE) is proposed. Also proposed is a parallel algorithm, well suited for implementation on Graphics Processing Units, for efficient computation of the shortest geodesic paths. Another low level geometric descriptor, termed as the Biharmonic Density Estimate, is proposed to provide an intrinsic geometric scale space signature for multiscale surface feature-based representation of deformable 3D shapes.

The computer vision and graphics communities rely on mid-level geometric understanding as well to analyze a scene. Symmetry detection and partial shape matching play an important role as mid-level cues. A comprehensive framework for detection and characterization of partial intrinsic symmetry over 3D shapes is proposed. To identify prominent overlapping symmetric regions, the proposed framework is decoupled into Correspondence Space Voting followed by Transformation Space Mapping procedure. Moreover, a novel multi-criteria

optimization framework for matching of partially visible shapes in multiple images using joint geometric embedding is also proposed.

The ultimate goal of geometric shape analysis is to resolve high level applications of modern world. This dissertation has focused on three different application scenarios. In the first scenario, a novel approach for the analysis of the non-rigid Left Ventricular (LV) endocardial surface from Multi-Detector CT images, using a generalized isometry-invariant Bag-of-Features (BoF) descriptor, is proposed and implemented. In the second scenario, the geometric regularity and variability of the cortical surface fold patterns at the 358 Dense Individualized and Common Connectivity-based Cortical Landmarks (DICCCOL) sites are quantitatively analyzed using Geodesic Context Histogram, a histogram constructed using GFE values in the spatial neighborhood of a surface point. In the third and final application scenario, we formulate a partial shape matching based technique, that can analyze the structure of the geometric shapes of the images and match them in a concise and meaningful manner directly, rather than relying on metadata to solve the problem of Content Based Image Retrieval.

INDEX WORDS: Shape Analysis, Geometric Features, Partial Shape Matching, Joint Geometric Embedding, Geodesic Field Estimate, Biharmonic Density Estimate, Cardiac Morphology, Cortical Folding Patterns

Geometric Analysis of Shapes and Its application to Medical Image Analysis

by

Anirban Mukhopadhyay

B.Tech., West Bengal University of Technology, 2009

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2014

Geometric Analysis of Shapes and Its application to Medical Image Analysis

by

Anirban Mukhopadhyay

Approved:

Major Professor: Suchendra M. Bhandarkar

Committee: Tianming Liu
Khaled Rasheed

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
University of Georgia
August 2014

Dedication

To the person I look up to for my entire life, who amongst other numerous things, gave me the wings of mathematical realizations from potato selling to proving theorems and literary realizations from Pandab Goenda to Aranyak, my Jethu.

Geometric Analysis of Shapes and Its application to Medical Image Analysis

Anirban Mukhopadhyay

July 26, 2014

Acknowledgments

First of all, my sincere thanks go to my major advisor, Professor Suchendra Bhandarkar for his excellent guidance and continuous support which played the most significant role in shaping my dissertation. Dr. Bhandarkar is single handedly responsible of taking me out of my comfort zone every single day to acquire knowledge from anywhere and everywhere. His unparalleled dedication to constantly reshape my work is an inspiration, I especially remember those times before submission deadlines, when he'd work on my papers and return them even at 2 am in the morning. I am grateful to Professor Tianming Liu for his brilliant insights into various aspects of medical image analysis and computer vision in general. Moreover, his financial support enabled me to attend premier research conferences which in turn gave me the opportunity of presenting to and interacting with internationally acclaimed scientists. My interaction with Professor Khaled Rasheed also had a significant impact on the machine learning sections of my dissertation. I'd also like to mention the meetings, phone calls and numerous e-mail interactions with Dr. Zhen Qian of Piedmont Heart Institute, Atlanta, which significantly enhanced my understanding of the clinical aspects of medical imaging research.

I'd like to express my gratitude to Dr. Fatih Porikli and Dr. Jeroen Van Baar of Mitsubishi Electrical Research Labs as well as Dr. Yefeng Zheng of Siemens Corporate Research for providing me the unique and rewarding opportunity of working at two extremely prestigious industrial research labs. The environment in the Visual and Parallel Computing

Laboratory (VPCL), directed by Professor Bhandarkar was ideal to develop my research skills. Especially, my dissertation is heavily dependent on three of my student collaborators, Austin New, who was a B.S. student and worked all the way to SIGGRAPH 2012, Chul Woo Lim, who was a M.S. student and developed single handedly the CUDA implementation of all-pair geodesic path determination algorithm and Arun CS Kumar, who is currently a PhD student and pushing my boundaries every single day by asking me challenging questions. The technical discussions with my past and present labmates i.e. Somenath Das, Karan Sharma, Kyle Krafka, Naveen Aitha, Hari Devulapally and Sowmya Shivakumar were also very interesting. In addition, I sincerely acknowledge the help of technical and administrative support staff of our department which include Nathan Steward, Piotr Mistzal, Ken Powell, Tammy Bridges, Suzy Sanders and Jean Power.

During my formative years, I got guidance and support from Professor Sandip Sarkar, Saha Institute of Nuclear Physics (SINP). I consider my days and experience at the SINP as the greatest inspiration for choosing scientific research as a career and Dr. Sarkar, along with Safi-da and Subhajit-da changed my whole outlook towards research. Amit-da's guidance and genius in all aspects of Kolkata based education was also an eye-opener. A Ph.D. degree program does not provide many social interaction opportunities, but the tight knit group of my fellow PhD students and their families Mayukh, Adrijo and Bhargabi, Arya and Dipannita and of course Hildi, made Athens home away from home. I dearly remember those FIFA nights, heated movie discussions and leg-pulling for no apparent reason. I'd also like to acknowledge Ashis-da, Bishu-da, Priyo-da and Shameem-da, who helped to settle this bong newcomer in a foreign land in the very initial days of his PhD. I also fondly recall the long phone conversations and facebook chats with Kaushik, Saikat, Arindam and Evan.

My doctoral research would not have been possible without my loving and supporting family. Jethu, Ma, Baba, Pishi, Borojethu, Jethi, Muia and of course Kuttush - I can't express enough in words, how much I am grateful to all of them; their wishes are the main

driving force for me. And last but not the least, a very special thanks to my girlfriend Srijita for designing my research posters, supporting me in the most difficult of times and most importantly, tolerating me in all the nitty-gritty details of life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the book it said: "Boa constrictors swallow their prey whole, without chewing it. After that they are not able to move, and they sleep through the six months that they need for digestion."

I pondered deeply, then, over the adventures of the jungle. And after some work with a colored pencil I succeeded in making my first drawing. My Drawing Number One ...

I showed my masterpiece to the grown-ups, and asked them whether the drawing frightened them.

But they answered: "Frighten? Why should any one be frightened by a hat?"

My drawing was not a picture of a hat. It was a picture of a boa constrictor digesting an elephant. But since the grown-ups were not able to understand it, I made another drawing: I drew the inside of a boa constrictor, so that the grown-ups could see it clearly. They always need to have things explained.

— Antoine de Saint-Exupry, *The Little Prince*

1.1 Motivation

This famous story is suggestive of the fact that even from early childhood, understanding and analyzing shapes play an important role towards perceiving and reconstructing the visual world. Our subconscious analyzes shapes with such ease in daily life, that most of the time we forget about the intricacies of the process. In the area of computer vision and graphics, however, imitating the human visual system to understand shape has proven to be a challenging task. In particular, the difficulty increases as soon as complex deformations are introduced. As a result, one key problem of computer vision and graphics is to develop techniques for analyzing the properties of shape and describing the deformations in a structured manner.

Shapes in 2D are often represented by contours or RGB images and in 3D by their bounding surfaces which are subsequently discretized as meshes or point clouds in various computer vision and graphics applications. Although these representations are suitable for visualization, they lack the descriptive qualities needed for computer vision and graphics applications such as shape comparison, shape matching, symmetry detection, fine grained recognition, medical image analysis to name a few. As a result, it is particularly important to derive low-level robust geometric features to uniquely describe points of a shape in order to enable these applications. In this thesis, we have developed two low-level geometric feature description techniques i.e. Goedesic Field Estimate (GFE) and Biharmonic Density Estimate (BDE) for description.

Mid-level understanding of shapes is also very important for several computer vision and graphics applications such as partial shape matching, symmetry detection. Matching of partially visible shapes across multiple images that exhibit extreme variations in scale, orientation, viewpoint and illumination and also instances of occlusion is a core unsolved problem in computer vision. Solving the partial shape matching problem can potentially unlock many

more high-level computer vision problems. We approach partial shape matching as a multi-criteria optimization problem where the distance between two shapes is represented by a joint geometric embedding distance measure. Mid-level understanding of Symmetry is important because of its ubiquity in nature and also in artificial man-made objects. The detection and characterization of shape symmetry has attracted much attention in recent times, especially within the computer graphics community because of its wide applicability in higher-level problems ranging from medical imaging to architectural representation. Although most of the existing literature has focused on the detection of extrinsic symmetries, there has been a steadily growing interest in detection and characterization of intrinsic symmetries. We have approached the detection and characterization of overlapping intrinsic symmetry in two phases, Correspondence Space Voting (CSV) for symmetry initialization, followed by Transformation Space Mapping (TSM) for symmetry generalization and characterization.

The ultimate goal of computer vision is to analyze at a high-level, compelling applications relevant to modern society. To that end, low-level and mid-level image and shape understanding provides a solid foundation. In particular, we concentrated on three different high-level applications of geometric shape analysis.

The first application deals with the geometric description of the complex endocardial surface morphology. The limitations of conventional imaging techniques have hitherto precluded a thorough and formal investigation of the complex morphology of the left ventricular (LV) endocardial surface and its relation to the severity of Coronary Artery Disease (CAD). However, recent developments in high-resolution Multirow-Detector Computed Tomography (MDCT) scanner technology have enabled the imaging of the complex LV endocardial surface morphology in a single heart beat. Analysis of high-resolution Computed Tomography (CT) images from a 320-MDCT scanner allows for the non-invasive study of the relationship between the percent diameter stenosis (DS) values of the major coronary arteries and localization of the cardiac segments affected by coronary arterial stenosis. In the proposed

technique, the endocardial surface morphology is formally quantified using a 3D Bag-of-Features model.

The second application deals with the analysis of regularity of cortical surface folding patterns. Owing to the increasing availability of Diffusion Tensor Imaging (DTI) data, there is a growing interest in assessing the structural differences in neural connectivity within cortical networks in diseased brains and healthy controls. The quantitative assessment of the geometric regularity and variability of the cortical surface has been hindered thus far due to the difficulty in formalizing a representation of the cortical surface folding patterns and establishing their correspondence across individual brains. We proposed a GFE-based scheme to formalize the regularity and variability across cortical surface folding patterns.

For the third and final application, we focus on the field of content based image retrieval (CBIR). CBIR is facing new challenges due to the rapid growth of available image data, thanks to photo sharing sites such as Flickr and Picasa. Whereas traditional audiovisual archives come with carefully curated metadata, allowing easy access based on a predefined thesaurus, user-generated images are hardly ever annotated apart from a few, and often with not very informative tags. The same holds for many older image archives that only recently have been digitized. This calls for methods that can analyze the images directly rather than relying on the attached metadata. In particular, we show how the proposed CBIR technique equipped with the proposed mid-level partial shape matching technique can effectively tackle the problem of widely varying imaging and viewing conditions.

1.2 Contributions of the Dissertation

1.2.1 Low-level Shape Analysis

1. An intrinsic geometric surface signature based on the GFE, termed as the GFE contextual histogram (GCH), is proposed.

2. A parallel version of the all pairs geodesic path determination algorithm, a key element in the GFE computation, is designed and implemented using GPUs to ensure that the GFE computation is indeed scalable for large datasets

3. To the best of our knowledge, the proposed BDE is one of few works that formulates a multiscale 3D shape signature as a distance density function, that can effectively capture both, the global shape and local surface detail, depending on the scale parameter value.

4. The BDE is one of the first surface descriptors where the intrinsic geometric scale space parameter can be directly related to the neighborhood size r. This makes the BDE both, intuitive and easy to relate to the perceivable 3D surface geometry.

1.2.2 Mid-level Shape Analysis

1. The proposed symmetry detection technique employs the functional map representation in conjunction with the correspondence space voting technique.

2. Robust and meaningful characterization of the symmetry transformation via formulation of a symmetry space which allows one to quantitatively distinguish between instances of simple and complex intrinsic symmetry. To the best of our knowledge, this quantitative distinction has not been attempted in the published literature.

3. Enabling recovery of the symmetry groups via clustering on the functional maps.

4. The formulation of a novel multicriteria optimization framework to address the hitherto unsolved problem of partial shape matching across images where the variations in imaging and viewing parameters are truly challenging.

5. The introduction of a new benchmark dataset wherein the variability ranges over several imaging and viewing parameter such as illumination (day versus night), viewpoint,

age of structures (historic versus new), presence of occlusion, presence of partially constructed structures, and inclusion of sketches and/or paintings of objects along with their captured images.

1.2.3 High-Level Shape Analysis

1. The Bag-of-Features (BoF) framework for non-rigid shape analysis is adapted for the purpose of cardiac shape analysis which is an important problem in cardiac imaging in particular and medical imaging in general.

2. A geometric and machine learning-based model of the relationship between localized changes in the LV endocardial surface morphology and the incidence and extent of stenosis in specific coronary arteries is proposed. To the best of our knowledge, this is one of the first attempts to model this complex clinical relationship in a mathematically structured manner.

3. The formulation of the partial shape matching framework for solving the challenging computer vision problem of content-based image retrieval.

4. The characterization of geometric regularity and variability of the highly irregular cortical surface folding patterns at the DICCCOL sites using GCH.

1.3 Organization

Rest of the dissertation is divided into six different chapters:

Chapter 2: Geodesic Field Estimate and Analysis of Cortical Surface Folding Patterns

Chapter 3: Biharmonic Density Estimate

Chapter 4: Intrinsic Symmetry Detection and Characterization

Chapter 5: A 3D BoF Descriptor for Analysis of the Left Ventricular Endocardial Surface Morphology

Chapter 6: Partial Shape Matching Using Graph Embedding and Its Application to Content-based Image Retrieval

Chapter 7: Discussion and Conclusions

Chapter 2

Geodesic Field Estimate and Analysis of Cortical Surface Folding patterns

2.1 Introduction

Owing to the increasing availability of Diffusion Tensor Imaging (DTI) data, there is growing interest in assessing the structural differences in neural connectivity within cortical networks in diseased brains and healthy controls [5]. However, the fundamental issue underlying structural connectivity assessment from DTI data is the localization of network nodes with geometrically meaningful cortical regions of interest (ROIs). In particular, the complex surface geometry of the brain, manifest in the cortical surface folding patterns, provides important cues for the prediction of cortical cytostructure and function [3] thereby suggesting the regularity of cortical surface folding patterns. On the other hand, many studies demonstrate the remarkable variability of the cortical surface folding patterns and their inherent complex geometry [14].

The quantitative assessment of the geometric regularity and variability of the cortical surface has been hindered thus far due to the difficulty in formalizing a representation of

the cortical surface folding patterns and establishing their correspondence across individual brains. Despite their paucity, such formal assessment studies are critical for several core unsolved problems in human brain mapping such as, brain image registration, brain image segmentation, and cortical shape analysis. For instance, formulation of effective morphological or connectional features for brain image registration and segmentation would be very challenging without prior knowledge of the geometric regularity and variability of cortical surface folding patterns. Moreover, such prior knowledge has been shown to be very useful for brain image registration [11] and could potentially benefit functional brain mapping via functional Magnetic Resonance Imaging (fMRI) signal extraction and activation detection.

More recently, a dense map of 358 cortical landmarks, termed as Dense Individualized Common Connectivity-based Cortical Landmarks (DICCCOLs) [26], has been identified and validated. Each DICCCOL site possesses white matter fiber connection patterns which are both consistent and predictive of the cortical functions of the corresponding site [26]. The high reproducibility and predictability of DICCCOL sites in individual brains have been demonstrated in recent studies using DTI data [26]. However, the regularity and variability of the 358 DICCCOL sites with respect to the cortical surface geometry is yet to be fully explored. This chapter focuses on the assessment of geometric regularity and variability of the cortical surface folding patterns at the 358 DICCCOL sites where the cortical surface is reconstructed as a triangular mesh from DTI data. A novel feature vector based on intrinsic surface geometry is employed to quantify the regularity and variability of the cortical surface geometry in the vicinity of each of the DICCCOL sites.

The Geodesic Field Estimate (GFE), a probability distribution of geodesic paths over a surface, has been shown to generate rich intrinsic geometric features of points on surface meshes [17], [17], as depicted in Figure 2.1. These intrinsic geometric features are used to construct contextual surface descriptors around each of the 358 DICCCOL sites. Extensive experiments are performed to validate the accuracy and robustness of the GFE on synthetic

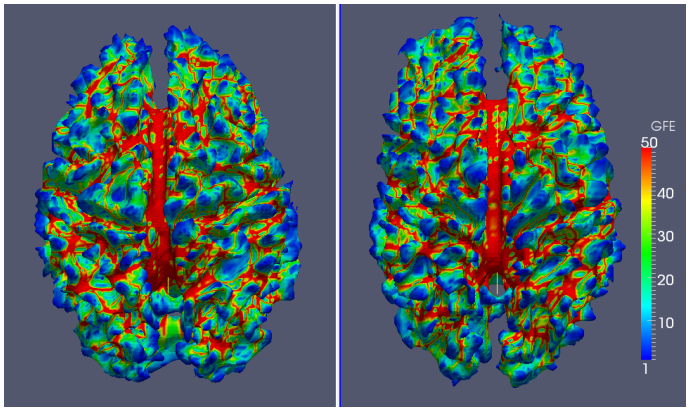

Figure 2.1: Visualization of the GFE on cortical surfaces of two different subjects.

data with folding patterns of incrementally increasing complexity and increasing levels of additive Gaussian noise. The value of the cumulative mean absolute deviation (MAD) for the contextual surface descriptor is computed at each DICCCOL site across different subjects and is considered as the measure of variability of the cortical surface folding pattern at that specific DICCCOL site. A major bottleneck for performing large-scale experiments with geodesic path-based surface descriptors is the computational complexity of geodesic path determination between all pairs of surface points. To address this computational bottleneck, a parallel version of the all-pairs geodesic path determination algorithm using Graphics Processing Units (GPUs) is proposed and shown to be broadly applicable to other medical imaging domains as well.

The results of experiments performed using DTI data from 31 healthy young adult brains demonstrate that some DICCCOL sites have significantly more regular cortical surface folding patterns than others. Overall, this study demonstrates the importance of geometric and morphological analysis of the complex cortical surface folding patterns which could be regarded as complementary to the analysis of white matter fiber connection patterns in the case of the more consistent DICCCOL sites. We envision that this study will offer novel insights

into MRI-based versus DTI-based brain mapping methodologies, where multimodal registration, mapping and analysis is performed using both, shape-based and connectivity-based features.

The primary contributions of this chapter are twofold. First, an intrinsic geometric surface signature based on the GFE, termed as the GFE contextual histogram (GCH), is proposed for the characterization of geometric regularity and variability of the highly irregular cortical surface folding patterns at the DICCCOL sites. Second, a parallel version of the all-pairs geodesic path determination algorithm, a key element in the GFE computation, is designed and implemented using GPUs to ensure that the GFE computation is indeed scalable for large datasets. The GPU code is made publicly available at `https://sites.google.com/site/geometricanirban/`.

The remainder of the chapter is organized as follows: Section 4.2 discusses the related work. Section 2.3 presents the theoretical framework underlying the GFE-based surface descriptor. Issues pertaining to the parallelization and GPU implementation of the GFE computation are described in Section 2.4. Section 2.5 describes the data preparation procedure whereas Section 3.5 presents the results of experimental validation. Finally, Section 6.6 concludes the chapter while outlining directions for future work.

2.2 Related Work

The complexity of the surface folding patterns of the human cerebral cortex demands study at various scales, from a local neighborhood of a cortical landmark to the entire cortical surface. Traditionally, the local analysis of cortical surface folding patterns is based on computation of the local surface curvature whereas analysis of the folding pattern of the entire cortical surface or lobe of the human brain is based on computation of the Gyrification Index (GI) [27] or spherical wavelets [24]. Techniques for analysis of cortical surface folding patterns

using surface descriptors have experienced a recent surge of interest. Toro et al. [22] have proposed using the surface ratio thereby extending the description from a global scale, such as one obtained using the GI, to a local scale. Zhang et al. [25] have proposed a parametric representation of cortical surface folding patterns with strong local shape representation capability.

It is important to note that all the works mentioned above either do not exploit the intrinsic surface geometry and at best use very simple extrinsic geometric surface descriptors (such as ones based on local surface curvature). Rich, comprehensive descriptors based on intrinsic surface geometry, such as the Wave Kernel Signature [AUBRY11], Heat Kernel Signature [25] and Discrete Surface Ricci Flow [30] have been widely used for shape representation and shape analysis in the computer graphics and computer vision communities but mostly for *simple* surfaces. In this chapter, we propose an intrinsic geometry-based surface descriptor, i.e., the GFE, to bridge the gap between the fields of geometric shape analysis and biomedical image understanding via analysis of *complex* and possibly *noisy* brain cortical surface data. The goal of the study is to enable formal quantitative assessment of the regularity and variability of the brain cortical surface folding patterns with respect to structural neural connectivity.

2.3 Theoretical Derivation of the GFE

The shortest distance between two points on a complete connected Riemannian manifold is defined by the length of the shortest geodesic path, i.e., the geodesic distance, between them. The geodesic distance is known to encode useful information about the underlying surface. However, determination of the geodesic distance is notoriously sensitive to small surface perturbations, making it difficult to use for robust shape analysis. In particular, for a complex and possibly noisy surface such as the cortical surface of the human brain,

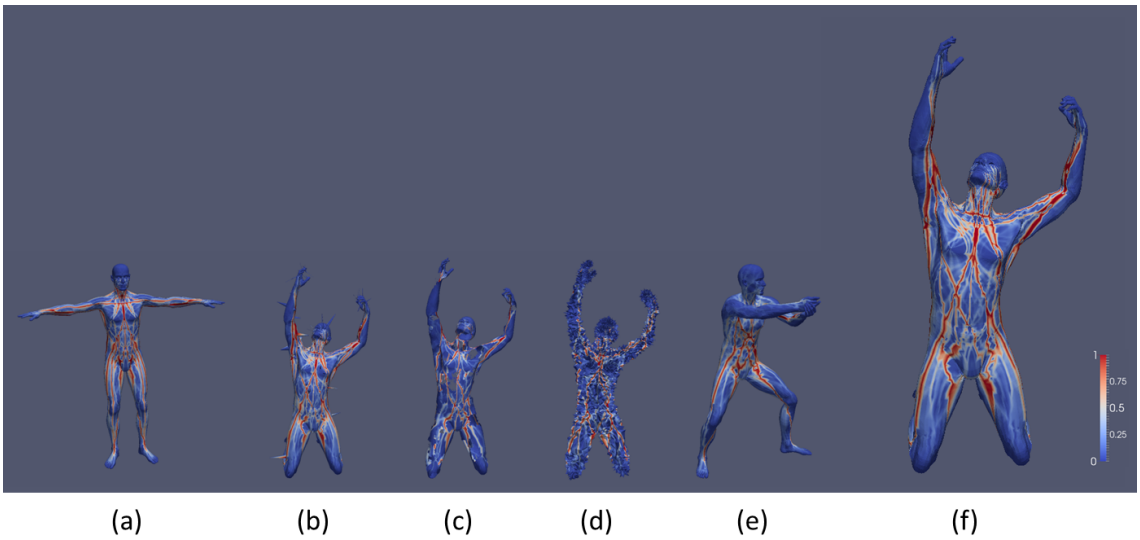

Figure 2.2: Robustness of the GFE in the face of challenging variations of the 3D *Human* shape model: (a) base case, (b) addition of shot noise, (c) changes in topology, (d) addition of Gaussian noise, (e) isometry and sampling and (f) changes in scale.

it is extremely challenging to extract meaningful information from the geodesic distance alone. The *shortest geodesic path*, on the other hand, conveys rich information about the underlying manifold. Since it is based on intrinsic surface geometry, the shortest geodesic path is invariant to the coordinate space in which the manifold is embedded. Moreover, we have observed that the overall structure of the geodesic paths is much more robust under different challenging real world conditions including surface perturbations [17]. To exploit the robustness property of geodesic paths, we have proposed a novel surface descriptor, termed as the *Geodesic Field Estimate* (GFE) [17].

The GFE for any point x on surface S can be defined as the probability of the shortest geodesic path between any two surface points p and q passing through x, i.e.,

$$GFE(x) = \text{prob}(x \in GP(p, q)) \tag{2.1}$$

where, for points $x, p, q \in S$, $GP(p, q)$ is the shortest geodesic path between points p and q. The GFE value at each surface point is computed using the all-pairs shortest geodesic path determination algorithm [8]. The GFE has been shown to be a rich and stable surface descriptor that is well suited for robust shape analysis [17]. As depicted in Figure 2.2, the GFE values are observed to be fairly robust to various challenging variations of the 3D *Human* shape model including addition of noise, changes in topology and isometric deformations.

Intuitively, the GFE can be visualized as *ropes* threading the *valleys* of the shape as depicted in Figures 2.1 and 2.2. Theoretically, the GFE can be shown to be a special case of the more general *fuzzy geodesics* [21]. As a result, the GFE inherits the property of robustness to noise and surface perturbations from the fuzzy geodesics while being more concise and informative than the latter.

2.3.1 Stability of the GFE

Theoretically, it has been shown that the stability of fuzzy geodesics can be quantified in terms of the Gromov-Hausdorff (GH) distance between two Riemannian manifolds as the shape deforms [21]. The GH distance has been used to measure the extent of shape deformation [16]. It has been shown that two shapes with a small GH distance have a provably small difference in their respective fuzzy geodesics which ensures that their corresponding GFE values are similar [21]. In particular, a tight bound for the fuzzy geodesics in terms of noise has been provided which also holds true for the special case of the GFE [21]. Moreover, to make the GFE robust to noise, we normalize the GFE with respect to the area of the triangular element when using a triangulated mesh-based representation of the underlying brain cortical surface in our implementation. The performance of the GFE in presence of noise is experimentally validated in Section 3.5 of this chapter.

2.3.2 Construction of the GFE-based feature vector

A GFE-based feature vector in the form of a GFE contextual histogram (GCH) is constructed for each DICCCOL site for each subject. A local neighborhood (i.e., ROI) for each DICCCOL site, comprising of the geodesically closest n surface points, is constructed. A k-bin histogram of GFE values (i.e, the GCH) is generated for the n-point ROI and represented as a k-tuple feature vector. The GCH feature vector at DICCCOL site x is formally denoted by $GCH_n^k(x)$.

Figure 2.3 shows similar GCH feature values across 5 subjects for a relatively regular DICCCOL site ROI #234 and very dissimilar GCH feature values across the same 5 subjects for a relatively irregular DICCCOL site ROI #80 where $k = 10$ and $n = 50$. The terms *GCH* and *GFE signature* are used interchangeably in the remainder of the chapter. The main reason behind using the contextual description of GFE is to describe the the coarse distribution of the GFE across the rest of the shape with respect to the given ROI.

2.4 Algorithm Description and Implementation

The high computational complexity of the all-pairs shortest geodesic path determination algorithm renders the use of the GFE highly impractical for most medical image analysis problems. A typical graph generated from the brain cortical surface mesh has a large number of nodes or vertices (in the range 40,000 - 50,000) but a relatively small number of edges (in the range 240,000 - 300,000). Also, the corresponding shortest-path search trees (where the root is the source node) are very deep and narrow since each node's connections are limited to its local neighbors with no shortcuts are available to reach farther nodes. It should be noted that the all-pairs shortest geodesic path determination procedure is the most computationally intensive aspect of the serial GFE algorithm. Figure 2.4(a) provides an outline of the serial GFE algorithm which employs Johnson's algorithm [8] for determining the shortest geodesic path between all surface point pairs.

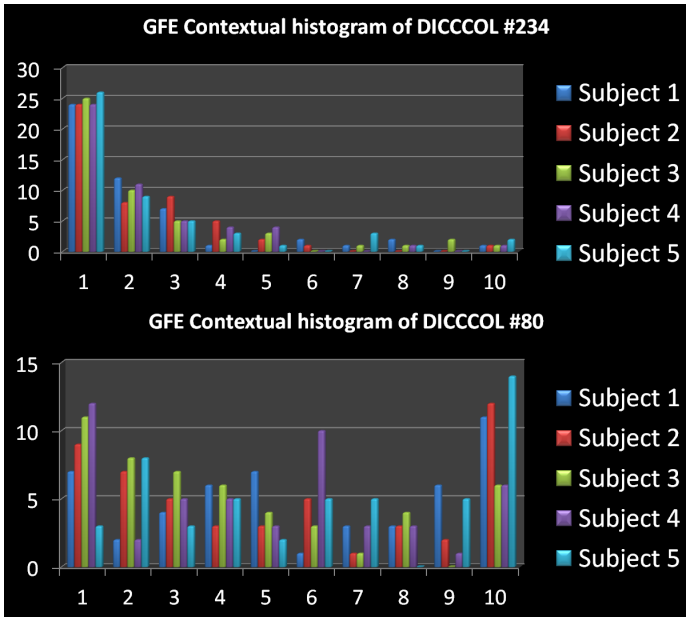

Figure 2.3: Visualization of the GFE context histogram at two DICCCOL sites for 5 different subjects.

We have proposed and implemented a novel instant update-propagate algorithm for parallelizing the all-pairs shortest geodesic path computation that is optimized for deep and narrow search trees. Figure 2.4(b) provides an outline of the GPU-optimized parallel implementation of the all-pairs shortest geodesic path algorithm along with a description of the CUDA kernel and the kernel calling procedure. Figure 2.5 depicts an execution instance of the instant update-propagate procedure on four streaming multiprocessors, each executing one block at a time. As can be noted, the updated distance cost is instantly reused by the subsequent block computations.

The proposed parallel algorithm is designed to utilize the GPUs more efficiently by letting the search procedure propagate to multiple levels in the search tree before global synchronization. A task parallel scheme [19] is adopted for computing multiple search trees originating from N different source points simultaneously, thus allowing for more efficient

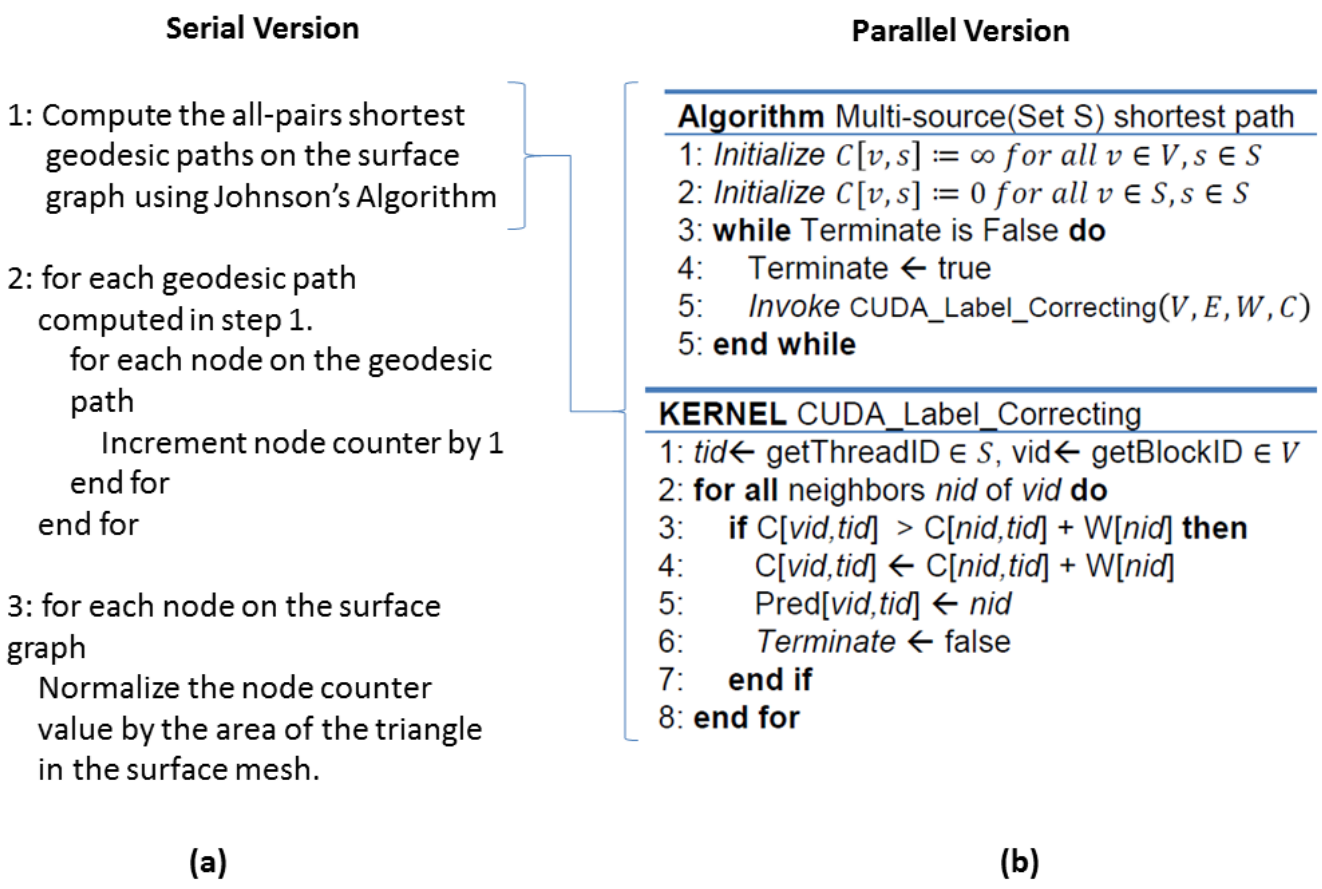

Figure 2.4: (a) Serial implementation of the GFE algorithm. Determining the shortest geodesic path between all pairs of surface points [8] is the most computationally intensive procedure within the serial GFE algorithm. (b) The GPU-optimized parallel implementation of the all-pairs shortest geodesic path algorithm. The CUDA kernel and the kernel calling procedure are described. Here $C[\cdot,\cdot]$ represents the cost of moving from one vertex to another, and *vid*, *tid* and *nid* respectively represent the source vetex, target vertex and the vertex through which the shortest path is traversing.

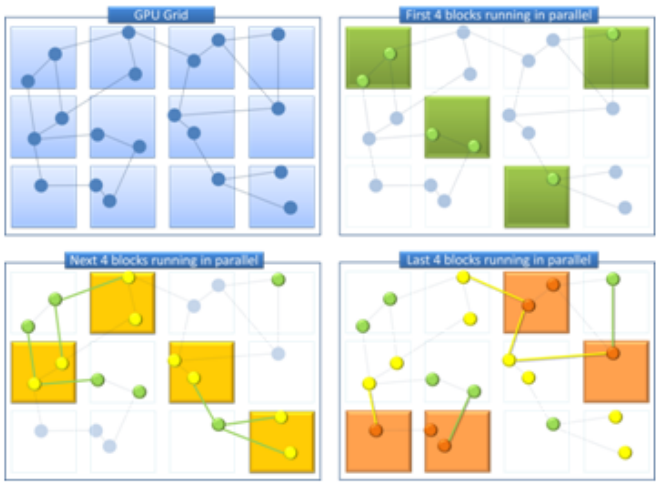

Figure 2.5: Execution example of the instant update/propagate strategy for computing the all-pairs shortest path problem on a GPU. The updated distance cost is instantly reused by the subsequent block computations. This example assumes there are four streaming multiprocessors, each executing one block at a time.

memory access patterns that exploit the high data throughput capacity of the GPU memory architecture. As depicted in Figure 2.6, the graph edge data are stored in the on-chip shared memory whereas the shortest distance label data is stored in global memory enabling greater memory access throughput. The GPU-optimized parallel code is made publicly available at `https://sites.google.com/site/geometricanirban/`.

2.5 Data Preparation

DTI data from 31 young adults from a publicly available database [23] are used in this study. The DICCCOL sites and connectomes identified and constructed from these DTI data [26] are regarded as the ground truth. The DTI data preprocessing is performed using the FSL software suite [6] which includes eddy current correction, skull removal, computing the Fractional Anisotropy (FA) image, and tissue segmentation. The cortical surface is

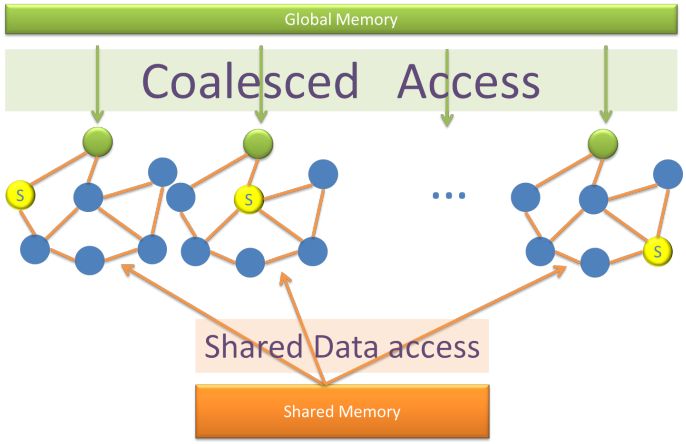

Figure 2.6: Memory access patterns for exploiting the high data throughput capacity of the GPU memory architecture. Instances of the shortest geodesic path computation with different starting points (highlighted in yellow) are processed in groups to leverage the GPU memory architecture. The graph edge data are stored in the on-chip shared memory whereas the shortest distance label data is stored in global memory.

reconstructed using the segmented FA image followed by fiber tracking performed using the MedINRIA software [LIU07]. The DICCCOL sites and connectomes are obtained from the preprocessed data using publicly available programs at `http://dicccol.cs.uga.edu/`.

In the DICCCOL framework, all the cortical landmarks are defined and predicted using DTI data. Therefore, the mapping of DTI-derived DICCCOL sites onto the MNI/Talairach atlas image has to rely on MR image registration techniques. Given the 358 DICCCOL sites from ten template brains with the corresponding structural MR images, the DICCCOL sites in each DT image of the template brains are registered with the corresponding MR images and warped onto the MNI template using the FSL FLIRT software tool [7] since it was observed to perform better than the alternatives [10]. Since there is no ground truth data for evaluating the correspondence of the DICCCOL sites with the MNI atlas image, the performance of the image registration algorithm is assessed entirely in terms of consistency resulting in a slightly higher accuracy for FSL FLIRT (6.29 mm) when compared to the

alternatives [10].

2.6 Experimental Results

2.6.1 Experimental validation using synthetic deformation

To the best of our knowledge, this is one of the first attempts to study the problem of surface regularity and variability at cortical surface ROI sites. Due to the paucity of benchmark brain datasets with manual ground-truth labeling, it is not possible to experimentally validate the performance of the proposed GFE signature on real data. To resolve this issue, the GFE is evaluated on simple model surfaces where synthetic deformations of increasing complexity are added to pre-selected surface points to validate the uniqueness property of the GFE and its ability to capture deformations at specific surface points. The conformal (i.e., angle preserving) deformations are added using the spin transformation technique developed by Crane et al. [2]. In their spin transformation technique, Crane et al. [2] propose a quaternionic Dirac operator along with an integrability condition on the conformal deformations. Discretization of the integrability condition results in a sparse linear system that is easy to solve and can be used to efficiently edit surfaces by manipulating their curvature. In particular, in the case of a simple model cylindrical surface, we successfully generate complex deformations by systematically increasing the surface curvature values at select surface regions in six discrete levels. The corresponding changes in GFE are noted and depicted in Figure 2.7. It can be observed that the GFE values remain fairly stable in the unperturbed regions of the model cylindrical surface. Thus, the GFE is observed to successfully differentiate between the regularity and variability of the object surface.

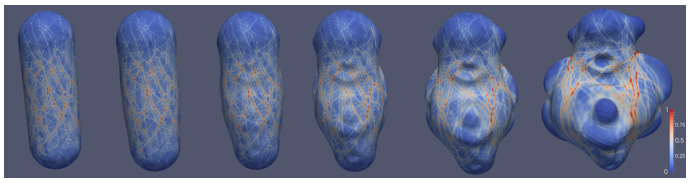

Figure 2.7: Visualization of regularity and uniqueness of the GFE on a simple model cylindrical surface with incremental deformations.

2.6.2 Parameter tuning

In this section, we describe the experimental tuning of the GFE signature parameters, i.e., the neighborhood size n and the number of bins k of the GCH, for optimal evaluation of the geometry of the human brain cortical surface. We rely on the synthetic deformations of the model cylindrical surface in order to tune the parameters n and k. The evaluation metric used in the tuning procedure is the difference in change in the GFE signature, denoted by $\Delta\epsilon$ where the change in GFE signature ϵ is computed for two different already known sets of points on the model surface, namely the set of points with maximal deformation and the set of points with minimal deformation computed using the methods described in [2]. The change in GFE signature ϵ is essentially the Euclidean distance, i.e., the L^2 norm, between the GFE signature of the original surface and that of the deformed surface normalized with respect to the GFE signature of the original surface. The value of $\Delta\epsilon$ is further normalized over the range of surface deformations. The parameters n and k that maximize $\Delta\epsilon$ are considered optimal. The $\Delta\epsilon$ evaluation metric is thus chosen to ensure that the resulting GFE signature exhibits low variability within the same class of deformations but high variability between different classes of deformations of the model surface.

We computed the value of $\Delta\epsilon$ for varying levels of deformations for different combinations of n and k as tabulated in Table 6.1. Since the combination of parameter values $k = 10$

Table 2.1: Tuning of parameters n and k for maximization of $\Delta\epsilon$ for the model cylindrical surface data.

k	n	$\Delta\epsilon$
10	30	0.33
10	50	0.43
10	100	0.32
20	30	0.32
20	50	0.42
20	100	0.38
30	30	0.35
30	50	0.42
30	100	0.37

and $n = 50$ was observed to maximize $\Delta\epsilon$, we used this set of parameter values for all the experiments reported in the remainder of the chapter.

2.6.3 Comparison with other descriptors

Since this is one of the first works to study the geometric regularity and variability of DICCCOL landmarks, we do not have any other existing methods against which to compare the accuracy of the proposed GFE signature. Instead, we have chosen to compare the performance of the proposed GFE signature with two state-of-the-art surface signatures most commonly encountered in the shape analysis literature for characterizing smooth triangular meshes in the context of our specific problem, i.e., the Heat Kernel Signature (HKS) [25] and the Wave Kernel Signature (WKS) [AUBRY11]. Specifically, we have employed the same synthetic data and performance metric described in Section 2.6.2 for this comparison. We have computed the value of $\Delta\epsilon$ for each of the surface signatures, i.e., HKS, WKS and GCH, for each incremental level of synthetic deformation of the model surface as shown in Table 2.2. It is interesting to note that the proposed GCH (with $k = 10$ and $n = 50$)

consistently outperforms both, the HKS and WKS for all levels of synthetic deformation of the model surface where the synthetic deformations are modeled to be similar in complexity to the variations of the brain cortical surface.

Table 2.2: Performance comparison of HKS, WKS and GCH (with $k = 10$ and $n = 50$) on the model cylindrical surface data for increasing levels of synthetic deformation. Higher values of $\Delta\epsilon$ indicate better performance.

Level	HKS	WKS	GCH
1	0.00	0.00	0.00
2	0.03	0.01	0.16
3	0.04	0.02	0.29
4	0.07	0.05	0.49
5	0.07	0.23	0.74
6	0.36	0.34	0.89
Mean	0.09	0.11	0.43

2.6.4 Robustness to noise

Extending the application of the GFE signature from the *simple* surfaces described in [17] to the less studied and less understood brain cortical surfaces, which could be potentially contaminated by noise, presents a significant challenge since it could dramatically alter the conclusions regarding the performance of the GFE signature. The immediate question here is whether the measured variability in the GFE and, in particular, whether the spatial separation of regular and irregular ROIs represents a characteristic of the brain or whether it is entirely a consequence of noise. In particular, there are three potential sources of noise that are present in the current brain cortical surface dataset:

1. Noise in surface geometry arising from noise in the MRI/DTI images or in the surface generation algorithm.

2. Noise in spatial localization of the DICCCOL sites.

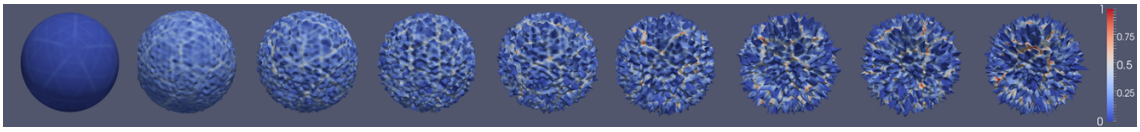

Figure 2.8: Robustness of GFE for different levels of synthetic white Gaussian noise added to the synthetic sphere model.

3. Spatial variation in any of the two noise sources mentioned above.

As shown previously [17], none of the aforementioned noise sources have been observed to hinder the successful computation of the GFE values. However, the aforementioned noise sources could potentially affect the successful clinical interpretation of the GFE results.

Although we have presented a qualitative argument for the robustness of the GFE signature to noise earlier in Section 2.3.1, we have designed an experimental validation procedure to demonstrate quantitatively the robustness of GCH with respect to noise. In particular, we assume that all the aforementioned sources of noise can be modeled as additive Gaussian noise. To simulate additive Gaussian noise, each vertex on the input model surface is randomly displaced along the local surface normal where the displacement is modeled by an independent and identically distributed (i.i.d.) random variable drawn from a zero-mean Gaussian distribution with a given value of σ (standard deviation). In particular, Figure 4.7 depicts the robustness of the GFE to different levels of synthetic Gaussian noise (characterized by the standard deviation σ) added to a synthetic spherical shape model. Note that the GFE values, as depicted in Figure 4.7, remain fairly consistent with increasing levels of additive Gaussian noise.

For each point on the synthetic shape, the L^2 norm between the values of the original GCH and the noisy GCH is computed and normalized with respect to the original GCH value to evaluate quantitatively the robustness of GCH to additive Gaussian noise. The

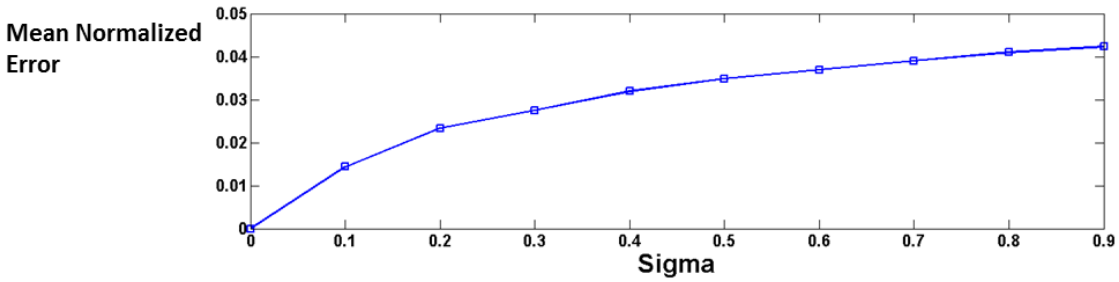

Figure 2.9: Robustness of the GCH for different levels of synthetic white Gaussian noise added to the synthetic sphere model.

mean normalized error in the GCH values is computed over all surface points for varying σ values (in the range $[0, 0.9]$) of the white Gaussian noise as depicted in Figure 2.9. It is interesting to note that the mean normalized error initially increases with increasing values of σ before settling upon a value of ≈ 0.035 for values of $\sigma > 0.4$, thus showing the robustness of proposed technique to additive white Gaussian noise.

2.6.5 Speedup results

The parallel GFE computation on a GPU was observed to achieve a speedup of 14 over its optimized CPU implementation (Johnson's algorithm [8] in the Boost C++ library `http://www.boost.org/`), taking less than a minute for each subject. The GFE computation was performed on a PC workstation with an NVIDIA GTX 480 GPU and an Intel Core i5-2400 CPU clocked at 3.4 GHz. Figure 2.10 demonstrates the speedup resulting from the GPU-optimized GFE computation on surface mesh graphs for the model shapes *horse*, *dog* and *human* from the SHREC 2010 dataset [15]. It can be observed that the proposed parallel algorithm for GFE computation yields speedup figures in the range $[24, 31]$ for the aforementioned model shapes. It should also be noted that these speedup figures are substantially higher than those obtained using the parallel algorithms described in [4]

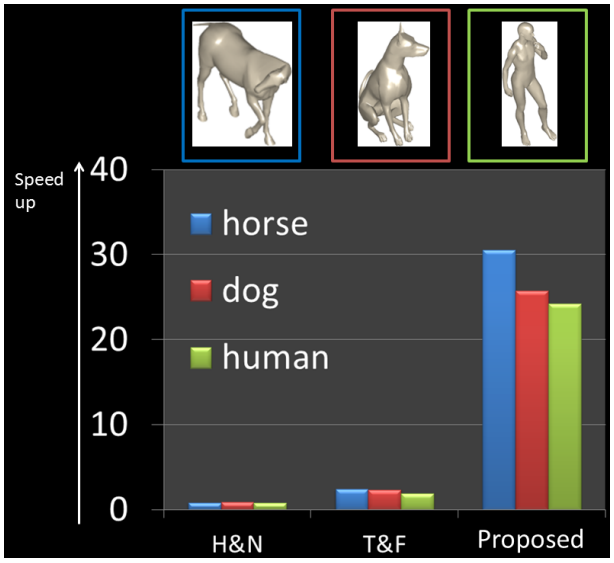

Figure 2.10: Visualization of the speedup resulting from GPU-based optimization. The proposed algorithm typically shows a 25-30 times speedup on the SHREC 2010 [15] dataset meshes with approximately 30000 - 50000 vertices. Here, H&N: [4], T&F: [19]

and [19].

2.6.6 Cumulative mean absolute deviation for measuring variability

The similarities of the GCH feature values for 358 DICCCOL ROI sites across 31 subjects are quantified by the cumulative mean absolute deviation (MAD) and shown as blue curves in Figure 2.11. It is evident that there is substantial variability across the ROIs in terms of the regularity/variability of their corresponding cortical folding patterns across the subject cohort. For instance, some ROIs, such as ROI #234 and ROI #94, show greater similarity in terms of surface geometry across the 31 subjects, whereas other ROIs, such as ROI #80 and ROI #252, exhibit greater variability across the same subject cohort. Based on the cumulative MAD values computed across 31 subjects, the top 5 percentile ROIs are considered

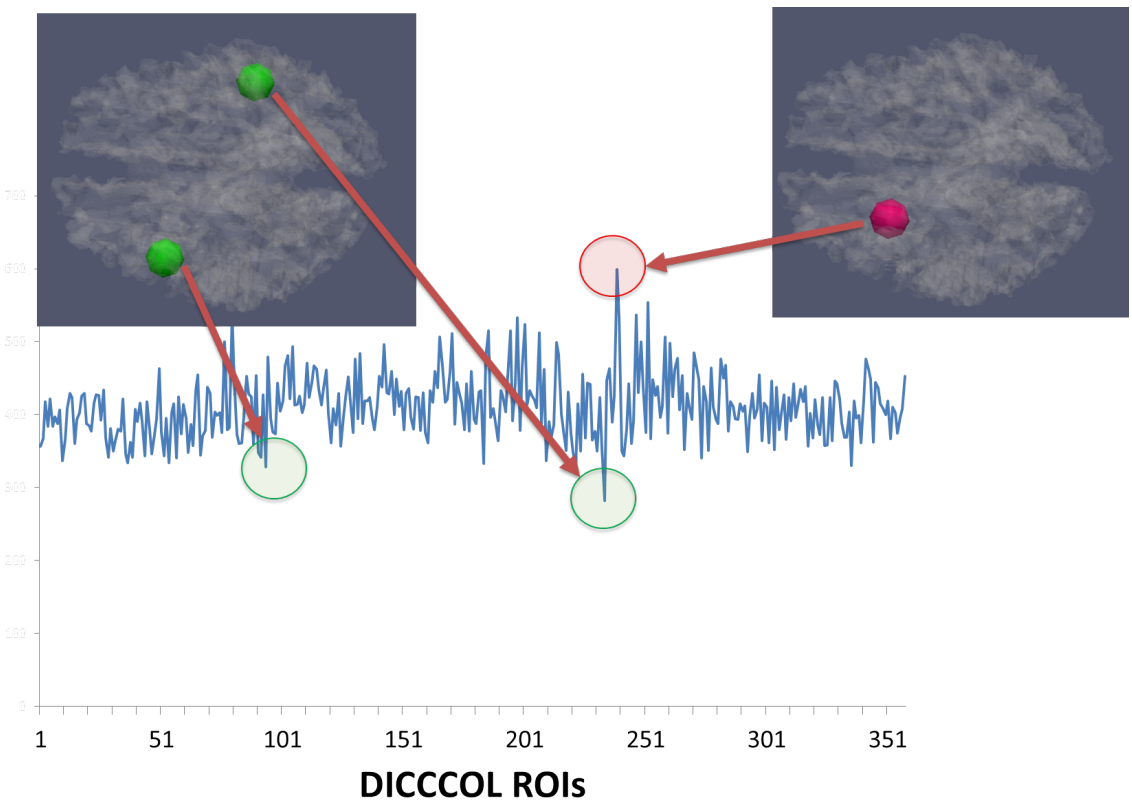

Figure 2.11: Plot of cumulative MAD values for the 358 DICCCOL ROI sites where lower MAD values denote higher regularity. Two DICCCOL ROI sites with higher regularity (green) and one DICCCOL ROI site with higher variability (red) are shown.

as the most variable and irregular across the subject cohort whereas bottom 5 percentile ROIs are considered the most regular. These ROIs are plotted in Figure 2.12 which reveal an interesting observation. The more geometrically stable ROIs are found towards the outer surface of the cerebral cortex whereas the least stable ones are found around the center. This phenomena can be justified by the fact that the DICCCOLs along the outer surface of the cerebral cortex are associated with various motor skills that are relatively uniform across the human subject cohort resulting in a more regular cortical folding patterns across subjects. Whereas, the DICCCOLs located near the center of the cerebral cortex are associated with more specialized skills such as linguistic skills which exhibit greater variation across the subject cohort, resulting in a more variable cortical folding patterns.

This phenomena calls for further and more thorough investigation.

2.7 Conclusions and Future Directions

A novel surface feature based on intrinsic geometry is proposed for analysis of cortical surface folding patterns at the DICCCOL sites in the human brain. Our study sheds new light on the relationship between the geometric regularity and structural regularity at DICCCOL sites within the cerebral cortex. Our study indicates that further research in morphological analysis of cortical surface folding patterns is needed. Specifically, the relative positions of the geometrically regular and geometrically variable DICCCOL sites within the cerebral cortex deserve more extensive and rigorous investigation. We plan to examine the possibility of using both, the cortical surface folding patterns and DTI-derived connectivity patterns to predict the locations of DICCCOL sites within individual brains, which could then be used for brain registration and mapping.

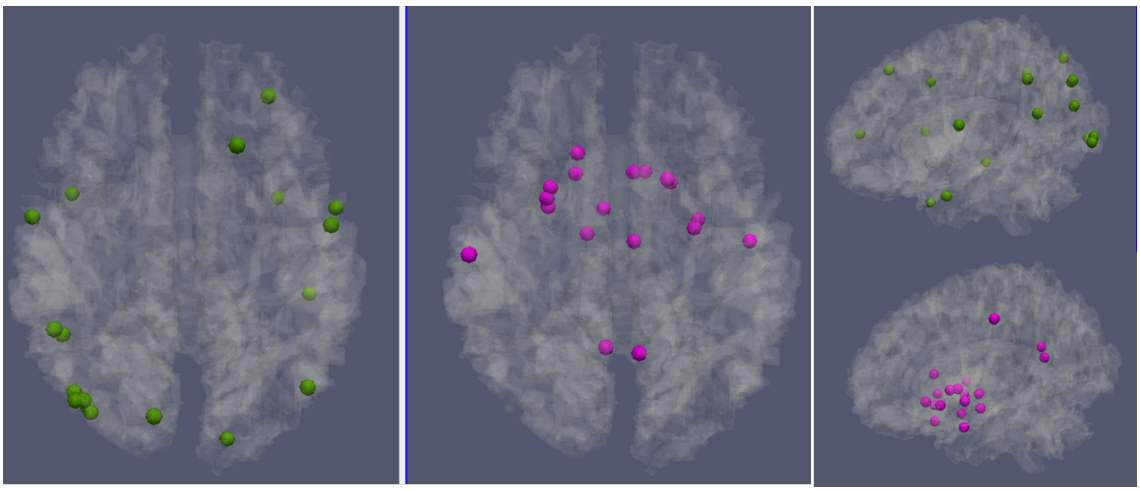

Figure 2.12: ROIs with cumulative MAD values in the bottom 5 percentile (regular) are plotted in green whereas ROIs with cumulative MAD values in the top 5 percentile (variable and irregular) are plotted in pink.

Bibliography

[1] Aubry, M., Schlickewei, U., & Cremers, D. The wave kernel signature: A quantum mechanical approach to shape analysis. In Proc. Computer Vision Workshops (ICCV Workshops), November 2011; pp. 1626-1633.

[2] Crane, K., Pinkall, U., & Schrder, P. Spin transformations of discrete surfaces. ACM Transactions on Graphics (TOG) 2011; 30(4).

[3] Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., et. al. Cortical folding patterns and predicting cytoarchitecture. Cerebral Cortex 2008; 18(8): 1973-1980.

[4] Harish, P. & Narayanan, P.J. Accelerating large graph algorithms on the GPU using CUDA. In Proc. Intl. Conf. on High Performance Computing (HiPC) 2007; 197208.

[5] He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., & Jiang, T. Regional coherence changes in the early stages of Alzheimers disease: a combined structural and resting-state functional MRI study. Neuroimage 2007; 35(2): 488-500.

[6] Jenkinson, M., & Smith, S. A global optimisation method for robust affine registration of brain images. Medical Image Analysis 2001; 5(2): 143-156.

[7] Jenkinson, M., Bannister, P., Brady, M., & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 2002; 17(2): 825-841.

[8] Johnson, D.B. Efficient algorithms for shortest paths in sparse networks. Journal of the ACM 1977; 24(1): 1-13.

[9] Li, K., Guo, L., Li, G., Nie, J., Faraco, C., Cui, G., et al. Gyral folding pattern analysis via surface profiling. NeuroImage 2010; 52(4): 1202-1214.

[10] Li, K., Guo, L., Zhu, D., Hu, X., Han, J., & Liu, T. Individual functional ROI optimization via maximization of group-wise consistency of structural and functional profiles. Neuroinformatics 2012; 10(3): 225-242.

[11] Liu, T., Shen, D., & Davatzikos, C. Deformable registration of cortical structures via hybrid volumetric and surface warping. NeuroImage 2004; 22(4): 1790-1801.

[12] Liu, T., Li, H., Wong, K., Tarokh, A., Guo, L., & Wong, S.T. Brain tissue segmentation based on DTI data. NeuroImage 2007; 38(1): 114-123.

[13] Liu, T., Nie, J., Tarokh, A., Guo, L., & Wong, S.T. Reconstruction of central cortical surface from brain MRI images: method and application. NeuroImage 2008; 40(3): 991-1002.

[14] Liu, T. A few thoughts on brain ROIs. Brain Imaging and Behavior 2011; 5(3): 189-202.

[15] Lian, Z., Godil, A., Fabry, T., Furuya, T., Hermans, J., Ohbuchi, R., et. al. SHREC10 track: non-rigid 3D shape retrieval. 3DOR 2010; pp. 101-108.

[16] Mmoli, F., & Sapiro, G. A theoretical and computational framework for isometry invariant recognition of point cloud data. Foundations of Computational Mathematics 2005; 5(3): 313-347.

[17] Mukhopadhyay A., Lim C. W., Bhandarkar S.M., Chen H., New A., Liu T., Rasheed K. & Taha T. Analysis of surface folding patterns of DICCCOLs using the GPU-optimized Geodesic Field Estimate. In Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI 2013) Workshop on Mesh Processing in Medical Image Analysis, 2013.

[18] New, A.T., Mukhopadhyay, A., Arabnia, H.R., & Bhandarkar, S.M. Non-rigid shape correspondence and description using geodesic field estimate distribution. In Proc. ACM SIGGRAPH 2012, pp. 96.

[19] Okuyama, T., Ino, F., & Hagihara, K. A task parallel algorithm for computing the costs of all-pairs shortest paths on the CUDA-compatible GPU. Intl. Jour. High Performance Computing and Networking 2012; 7(2): 87-98.

[20] Sun, J., Ovsjanikov, M., & Guibas, L. A concise and provably informative multiscale signature based on heat diffusion. Computer Graphics Forum 2009; 28(5): 1383-1392.

[21] Sun, J., Chen, X., & Funkhouser, T.A. Fuzzy geodesics and consistent sparse correspondences for deformable shapes. Computer Graphics Forum 2010; 29(5): 1535-1544.

[22] Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., & Paus, T. Brain size and folding of the human cerebral cortex. Cerebral Cortex 2008; 18(10): 2352-2357.

[23] Yan, C., Gong, G., Wang, J., Wang, D., Liu, D., Zhu, C., et. al. Sex-and brain sizerelated small-world structural cortical networks in young adults: a DTI tractography study. Cerebral Cortex 2011; 21(2): 449-458.

[24] Yu, P., Yeo, B. T., Grant, P. E., Fischl, B., & Golland, P. Cortical folding development study based on over-complete spherical wavelets. In Proc. Intl. Conf. Computer Vision (ICCV) 2007; pp. 1-8.

[25] Zhang, T., Guo, L., Li, G., Nie, J., & Liu, T. Parametric representation of cortical surface folding based on polynomials. In Medical Image Computing and Computer-Assisted Intervention 2009; pp. 184-191.

[26] Zhu, D., Li, K., Guo, L., Jiang, X., Zhang, T., Zhang, D., et. al. DICCCOL: dense individualized and common connectivity-based cortical landmarks. Cerebral Cortex 2013; 23(4): 786-800.

[27] Zilles, K., Armstrong, E., Schleicher, A., & Kretschmann, H.J. The human pattern of gyrification in the cerebral cortex. Anatomy and Embryology 1988; 179(2): 173-179.

[28] Zou, G., Hua, J., Lai, Z., Gu, X., & Dong, M. Intrinsic geometric scale space by shape diffusion. IEEE Transactions on Visualization and Computer Graphics 2009; 15(6): 1193-1200.

Chapter 3

Biharmonic Density Estimate

3.1 Introduction

Geometric shapes in 3D are often represented by their bounding surfaces which are subsequently discretized as meshes or point clouds in various computer graphics applications. Although mesh-based or point cloud-based representations are suitable for visualization, they lack the descriptive qualities needed for computer vision applications such as shape comparison, shape matching and shape symmetry detection. It is particularly important to derive robust features to uniquely describe 3D surface points in order to enable these computer vision applications. Scale space-based 3D surface and shape representations are often employed in computer vision applications because of their wide applicability and ability to describe the underlying structure at multiple scales of representation. The basic idea is to parameterize the surface or shape representation using a single family of scale parameters so that important spatial details are captured and characterized at multiple levels in the scale space hierarchy.

The specific goal of this chapter is to formulate a multiscale descriptor that allows for characterization of non-rigid 3D shapes in a manner such that point comparison, registra-

tion and description of the structural properties of the 3D shape can all be achieved in a concise and meaningful manner. In this chapter, we focus on non-rigid 3D shapes that are subject to isometric deformation. Under isometric deformation, the intrinsic geometry of the 3D shape remains invariant even though the extrinsic geometry may change. Hence, formal characterization of the intrinsic surface geometry is critical for understanding the underlying deformable 3D shape. Formulating an isometric deformation-invariant descriptor that captures the intrinsic geometry of the non-rigid 3D shape is, therefore, of particular importance.

Lipman *et al.* [14] introduced the *Biharmonic Distance Measure* (BDM) that possesses most of the desired properties for representation of 3D deformable shapes, i.e., the BDM is a metric, gradually varying, locally isotropic, globally shape-aware, isometric deformation-invariant, insensitive to noise and topology, practical to compute and parameter-free. The BDM has been shown to achieve a balance between the local and global properties of the underlying geometry via eigenvalue normalization [14]. oreover, the BDM is shown to bridge the gap between the *Geodesic Distance Measure* (GDM) [17] and the *Diffusion Distance Measure* (DDM) [25] (which models the heat diffusion process over a surface) for *small* values of the time parameter t, both of which capture *local* geometric information, and the DDM for *large* values for time parameter t, which effectively encapsulates the *global* properties of the underlying surface geometry. Although the DDM, with varying values for the time parameter t, has been shown to be useful for geometry processing problems [25], the time parameter t has no intuitive or obvious geometric interpretation.

In this chapter, we present the formulation of a novel shape descriptor based on the BDM, termed the *Biharmonic Density Estimate* (BDE) that essentially encapsulates the local density of biharmonic distances in the neighborhood of each 3D surface point. The degree of localization around each 3D surface point is dictated by a localization or scale parameter r, represented by a biharmonic distance sphere of radius r over the triangular

mesh neighborhood centered at that particular point. By varying the parameter r, one can construct a 1-parameter family of multiscale 3D shape signatures around each 3D surface point. As a result, the multiscale BDE descriptor is shown to be capable of describing the 3D shape as a whole, i.e., encapsulating the global 3D shape properties, or describing the local intrinsic surface, depending on the chosen value of the localization or scale parameter r.

In the proposed BDE signature-based representation, 3D shapes are generalized as points in a compact, connected 2-manifold that is equipped with a Riemannian metric. This results in a shape signature that encapsulates the intrinsic geometry of the 3D shape. Consequently, the intrinsic geometric scale space signature is invariant to how the manifold is actually perceived in 3D Euclidean space $\mathbb{R}^3$. In this chapter, the concept of intrinsic geometric scale space is proposed and formalized using the intrinsic geometric diffusion framework.

The proposed scale-aware BDE signature is shown to have a representation that is concise, efficient, informative and capable of encapsulating the intrinsic geometry of the 3D shape at multiple scales. The integration of the BDE with the proposed scale space framework is shown to result in a better understanding of the intrinsic geometry of the Riemannian 2-manifold and the underlying 3D shape. There are several applications in computer vision, such as shape comparison, shape matching, symmetry detection, to cite a few, that can potentially benefit from the proposed intrinsic geometric scale space framework. Our experimental results show that the multiscale BDE signature not only provides a rich description of local and global 3D surface structure, but is also unique, robust to sampling, invariant under isometric deformation, exhibits well explained scale space behavior and performs better than the state-of-the-art homogeneous feature detection techniques such as the Heat Kernel Signature (HKS) [25] and the Wave Kernel Signature (WKS) [AUBRY11]. The applications of the BDE in the detection of key components on a deformable 3D surface and determination of sparse point correspondences between two deformable 3D shapes are also

demonstrated.

The major contributions of the work presented in this chapter are two-fold. First, to the best of our knowledge, this is one of few works that formulates a multiscale 3D shape signature as a distance density function (which is distinct from its context descriptor counterpart [9]) that can effectively capture both, global shape and local surface detail, depending on the scale parameter value. Second, this is one of the first works where the intrinsic geometric scale space parameter is directly related to the neighborhood size r which is both, intuitive and easy to relate to the perceivable 3D surface geometry. This is in contrast to the scale parameters associated with the HKS and WKS where the relation between the scale parameter value and neighborhood size is not intuitive or obvious.

The remainder of the chapter is organized as follows: Section 3.2 discusses the related work, Section 3.3 presents the theoretical framework for the proposed BDE shape signature whereas the implementation issues are described in Section 3.4. Section 3.5 presents the results of our experimental validation whereas Section 3.6 demonstrates practical applications of the proposed BDE shape signature in the detection of key components on a deformable 3D surface and determination of sparse point correspondences between two deformable 3D shapes. Finally, Section 6.6 concludes the chapter while outlining directions for future work.

3.2 Related Work

The existing shape descriptors belong to two principal classes based on their degree of invariance. Shape Context [9], Shape Distribution [18] and Spin Images [9] are 3D shape descriptors that have been shown to be invariant under rigid-body transformation or deformation in $\mathbb{R}^3$. The Integral Invariants [15] and the Multiscale Local Signature [27] also belong to the class of 3D shape descriptors that are invariant to 3D rigid-body transformation. More recent research on shape description has focused on the formulation and

implementation of shape descriptors that are invariant to non-rigid deformation. Amongst the existing non-rigid deformation-invariant shape descriptors, the ones most related to the proposed BDE shape signature are based on the spectral properties of the Laplace-Beltrami operator designed for 3D surface characterization [AUBRY11], [25].

The classical computation of the discrete Laplacian on a mesh was introduced by Pinkall and Polthier [PINKALL93]. Taubin [26] employed the Laplacian for fair mesh surface design. Karni and Gotsman [10] used a spectral decomposition approach, based on computation of the Laplacian eigenvectors, for mesh compression. Levy [11] used the Laplace-Beltrami operator for geometry processing, followed by Reuter's influential work on Shape-DNA [22] and the work of Rustamov [23] who used the Laplace-Beltrami eigenfunctions to define an embedding of a 3D surface in the space of convergent sequences. The resulting formulation was shown to possess the desirable property of encapsulating all the intrinsic information for each point but was seen to be dependent on the choice of the eigenfunction basis. Even in the case of non-repeated eigenvalues, it was seen to be necessary to choose a sign for each basis component.

The problem of choice of the eigenfunction basis was resolved by the introduction of the Heat Kernel Signature (HKS) [25] and Wave Kernel Signature (WKS) [AUBRY11]. Although both, the HKS and WKS, are multiscale signatures, their scale parameters do not relate directly to the metric geometry of the local neighborhood of points on the 3D surface mesh. Note that there is no intuitive formulation connecting the time parameter (in the case of the HKS) or energy parameter (in the case of the WKS) to the parameters of the intrinsic sphere centered at the surface point of interest. The proposed BDE shape signature formulation has the advantage that the scale parameter exhibits direct association with the 3D geometry of the local neighborhood. An alternative formulation of intrinsic geometric scale space has been proposed by Zou *et al.* [30]. However, their formulation is based on Ricci flow shape diffusion which, unlike our formulation, is not directly related to the 3D

geometry of the local neighborhood and is also computationally much more expensive [30].

3.3 Theoretical Derivation

3.3.1 Biharmonic Distance Measure (BDM)

The *Biharmonic Distance Measure* (BDM) is similar in form to the *Diffusion Distance Measure* (DDM) [LING06]. The kernel of the BDM is based on the Green's function of the biharmonic differential equation. The Green's function is the impulse response of an inhomogeneous differential equation defined on a domain, with specified initial conditions or boundary conditions. In the continuous case, the (squared) biharmonic distance between two points x and y can be defined using the eigenvectors (ϕ_k) and eigenvalues (λ_k) of the Laplace-Beltrami operator [14] as follows:

$$d_B(x,y)^2 = \sum_{k=1}^{\infty} \frac{(\phi_k(x) - \phi_k(y))^2}{\lambda_k^2} \tag{3.1}$$

The above definition of the BDM is slightly different from that of the DDM where the denominator in the case of the DDM is $e^{2t\lambda_k}$. However, this subtle change ensures that the BDM formulation exercises greater control over the characterization of the global and local properties of the underlying manifold. Consequently, the BDM is fundamentally different from the DDM with significantly different properties.

The BDM, as expressed in equation (4.1), captures the rate of decay of the normalized eigenvalues λ_k of the Laplace- Beltrami operator; if the decay is too slow, it produces a logarithmic singularity along the diagonal of the Green's function [YEN07]. Alternatively, too fast a decay would basically ignore eigenvectors associated with higher frequencies, resulting in the BDM being too global in nature (i.e., the local surface details would be ignored). Lipman *et al.* [14], demonstrated that performing quadratic normalization provides a good

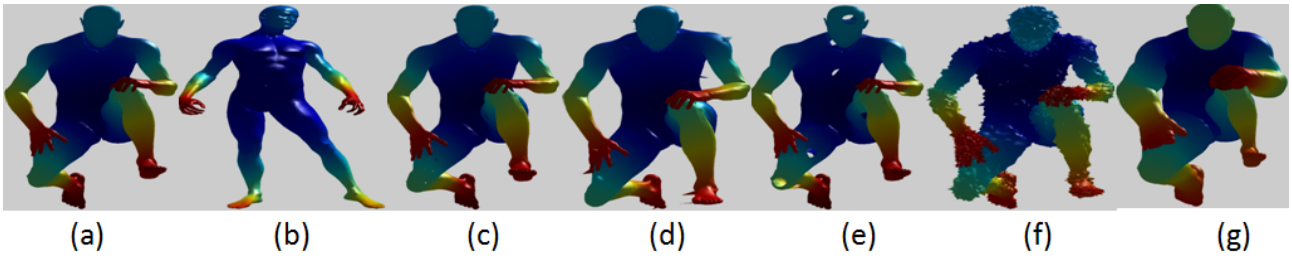

Figure 3.1: Robustness of the normalized Biharmonic Density Kernel under (a), (b) isometry, (c) micro-holes, (d) shot-noise, (e) holes, (f) noise and (g) local scaling. It is important to note the consistency of the BDK values across different deformations and surface perturbations.

balance, ensuring that the decay is slow enough to capture the local surface properties around the source point and yet rapid enough to encapsulate global shape information.

In particular, Lipman *et al.* [14] have theoretically proven two important properties of the BDM, i.e., that it is (i) a metric, and (ii) smooth everywhere except at the source point where it is continuous. The key observation is that for 3D surfaces, the eigenvalues λ_k, $k = 1, 2, \ldots$, of the Laplacian are an increasing function of k resulting in the continuity of the BDM everywhere and also smoothness everywhere except at the source point, where it has only a derivative discontinuity.

3.3.2 Biharmonic Density Kernel (BDK)

The *Biharmonic Density Kernel* (BDK) at a point x can be constructed from the eigenvectors (ϕ_k) and eigenvalues (λ_k) of the Laplace- Beltrami operator [14] as follows:

$$BDK(x, x) = \sum_{k=1}^{\infty} \frac{(\phi_k(x))^2}{\lambda_k^2} \tag{3.2}$$

The BDK provides a highly concise and informative representation of the underlying 3D mesh. The validity of the BDK is supported by the same information theoretic framework

proposed by Sun *et al.* [25] in case of the HKS. According to the theorem in [25], the diagonal elements of the kernel matrix are almost as informative as all the elements of the original kernel matrix. As a result, the BDK naturally inherits many of the intrinsic properties of the BDM. In particular, Figure 3.1 provides a qualitative assessment of the robustness and consistency of the BDK under different deformations and surface perturbations such as isometry, micro-holes, shot-noise, holes, noise and local scaling. The robustness and consistency of the BDK are experimentally verified in Section 3.5 of the chapter.

We argue in favor of the conciseness and richness of the information content of the BDK, by comparing it with the corresponding density functions of both the HKS and the WKS (3.2). Since the heat kernel is a collection of low-pass filters, the HKS descriptor is dominated by low frequencies that encapsulate the global structure of the underlying shape. Although this property enables the HKS descriptor to discriminate between distinct shapes, which usually differ greatly at coarse scales, it hinders the ability of the HKS descriptor to precisely localize the surface features.

Unlike the HKS that uses low-pass filters, the response of the WKS, in contrast, is band-pass. This reduces the influence of the low frequencies and allows better separation of frequency bands across the WKS descriptor dimensions resulting in superior feature localization. The WKS is well suited for point correspondence algorithms that rely primarily on local surface information but is unable to adequately represent the global shape information. The emphasis on precise localization of features allows matching algorithms based on the WKS to exhibit higher sensitivity (i.e., correct identification of true positives) whereas the emphasis on global features enables the HKS-based matching algorithms to exhibit higher specificity (i.e., correct identification of true negatives) [8]. The BDK on the other hand, has the unique property of being able to encapsulate both, local and global shape information [14] making it a very desirable candidate for the density kernel of a 3D shape signature. As seen in Figure 3.2, the extrema of the *Human* shape (e.g., the hands and the feet) do not

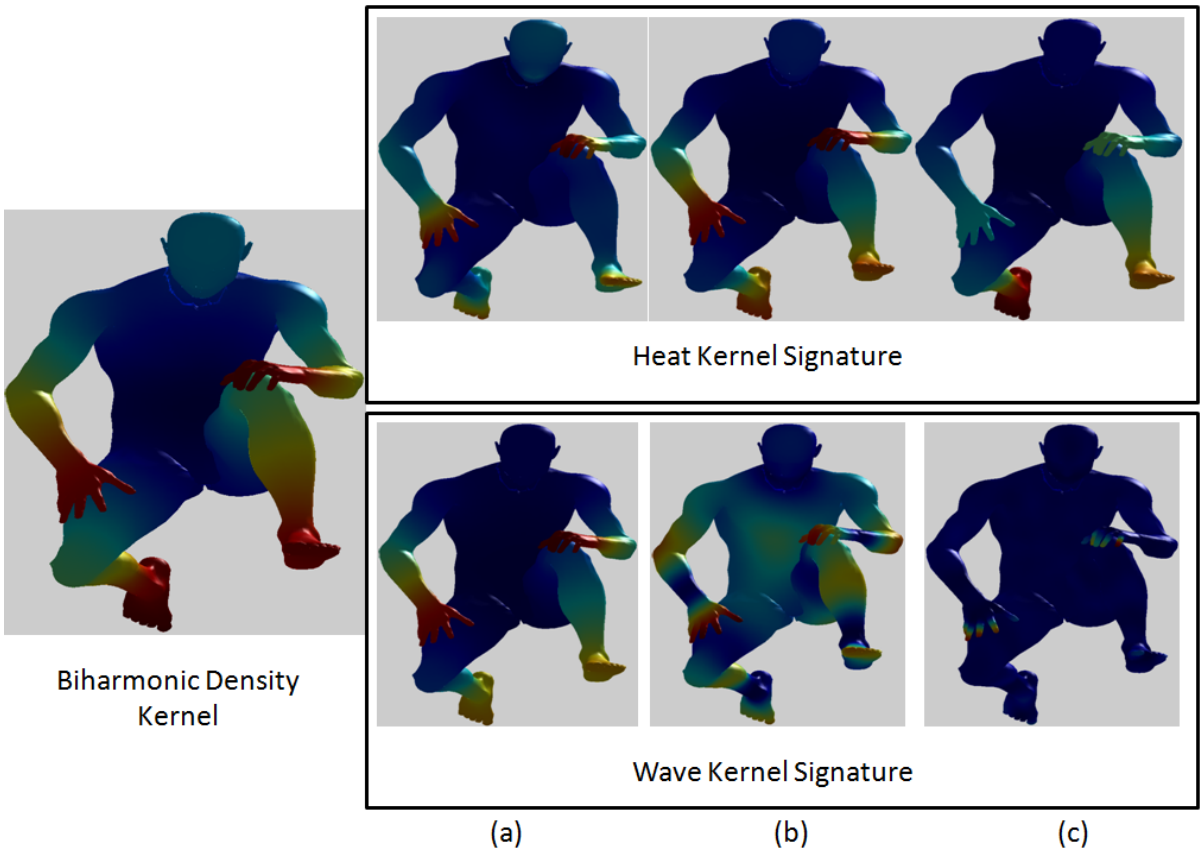

Figure 3.2: Comparison of the Biharmonic Density Kernel (BDK) with the Heat Kernel and the Wave Kernel computed at different diffusion scales (a) small, (b) medium and (c) large. Red indicates a high function value whereas, blue indicates a low function value and other colors represent intermediate values.

exhibit consistently high function values (indicated by the color red) over all values of scale in the case of the HKS and WKS. The BDK, on the other hand, simultaneously exhibits high values at the shape extrema and low values (indicated by the color blue) within uniform surface regions (e.g., the torso) of the *Human* model.

3.3.3 Biharmonic Density Estimate (BDE)

The *Biharmonic Density Estimate* (BDE), proposed in this chapter, is a novel 3D shape descriptor based on the BDK, that attempts to capture the local intrinsic geometric properties of the 3D surface in the neighborhood of a given surface point. In the continuous manifold, BDE can be formalized as the integral of $BDK(x, x)$ over the surface patch defined as the intersection of the biharmonic distance sphere of radius r with the surface, normalized by the surface area of the patch.

In the discrete case however, the BDE at a 3D surface point p computed over an r-neighborhood can be defined as:

$$BDE(p; r) = \sum_{x \in B(p;r)} W_r(x) \cdot BDK(x, x) \tag{3.3}$$

where $B(p; r)$ (i.e., the r-neighborhood) is the set of surface points contained within a biharmonic distance sphere of radius r defined over the triangular mesh neighborhood centered at point p and $W_r(x)$ denotes the weight associated with point x. In particular, we have experimented with two different averaging or weighting schemes: (i) simple averaging, i.e., $W_r(x) = \frac{1}{N_{B(p;r)}}$ where $N_{B(p;r)}$ denotes the cardinality of $B(p; r)$ and, (ii) Gaussian averaging where the weights $W_r(x)$ are derived from a zero-mean Gaussian distribution as follows:

$$W_r(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{d^2(p, x)}{2\sigma^2}\right) \tag{3.4}$$

where $d(p, x)$ denotes the biharmonic distance between points p and x and $\sigma = r/3$. Note that $r = 3\sigma$ ensures that over 99% of the weight of the Gaussian distribution is contained within $B(p; r)$.

Since the BDM has been proven to be a metric that is invariant to isometric deformation [14], the BDE can also be seen to satisfy the *isometric deformation-invariance* property. Additionally, the BDE is also able to capture the *intrinsic geometry* of the underlying 3D surface. Moreover, as a density estimate, the BDE is also able to successfully encode the local geometric properties of the 3D shape. Since the BDE is based on the distribution of the values of the BDK, it is very stable under surface perturbations.

3.4 Implementation

For a general compact manifold the BDM can be computed using the corresponding Laplace-Beltrami operator. Many schemes have been proposed to estimate the Laplace-Beltrami operator from discrete meshes [BELKIN08], [PINKALL93]. The cotangent scheme [27] is the most commonly used method in the computer graphics community for this purpose. However, it has been shown that the cotangent scheme does not converge in general [27], [28]. On the other hand, Belkin *et al.* [BELKIN08] have proposed a convergent scheme called the *mesh Laplace* operator. The convergence of this operator does not require the triangles in the mesh to be well-shaped like the ones generated via the Finite Element Mesh (FEM) methods associated with the cotangent scheme. For this reason, we have used the mesh Laplace operator to estimate the Laplace-Beltrami operator.

Given a mesh with n vertices, the mesh Laplace operator can be represented as L, a sparse matrix of size $n \times n$. Matrix L can be rewritten as $L = A^{-1}W$ where A is a positive diagonal matrix. The elements $a_{i,i}$ of A represent the area associated with vertex i, whereas W is a symmetric semi-definite matrix.The generalized eigenvalue problem $W.\phi = \lambda.A.\phi$ has

a full set of real eigenvalues and eigenvectors under these conditions. Having determined the eigenvalues and eigenvectors for the aforementioned generalized eigenvalue problem, the BDM between two points is computed followed by the computation of the BDE. The entire BDE computation is implemented on a PC workstation with a 2.4 GHz CPU and 24 GB RAM. The MATLAB mesh processing package [20] is used for visualization of all results.

3.5 Experimental Validation

Several experiments were performed to demonstrate the properties of the proposed BDE shape signature. The experiments were carried out using 3D models from the SHREC 2010 benchmark dataset [7], Non-rigid World dataset [5] and the TOSCA benchmark dataset [6].

3.5.1 Uniqueness Property

As shown in Figure 3.3, we note the BDE signatures of distinct points on the surface of the *Cat* model to demonstrate the uniqueness property of the BDE signature. It is particularly interesting to note that the BDE signatures differ significantly at finer scales (smaller r values). Also, the general shape of the BDE curve is a very interesting descriptor as it can serve as a good predictor of the overall surface shape. The two rear legs of the *Cat* model generate BDE signatures that are distinct from those of its front legs. It is also interesting to note that, for points on the legs, the BDE signature exhibits its first big hump as the expanding r-ring neighborhood (i.e., the biharmonic distance sphere) approaches the location where the leg is joined to the body of the *Cat* model. In contrast, the BDE signature of the point on the tail exhibits a completely different trend at finer and coarser scales.

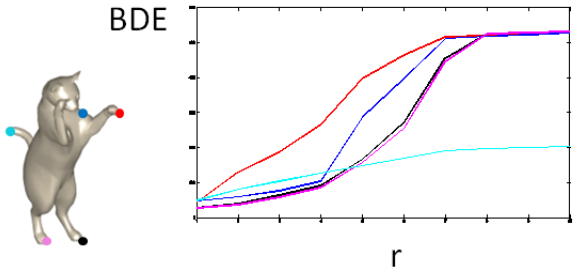

Figure 3.3: Uniqueness of BDE signature for different points selected on the *Cat* model. The surface points and their corresponding BDE signatures are color-coded.

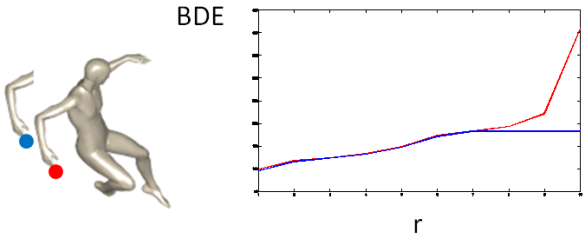

Figure 3.4: Invariance of the multiscale BDE signature of the *Human* model under occlusion.

3.5.2 Robustness to Occlusion

We performed a study to test the ability of BDE to describe adequately partially occluded surfaces. As shown in Figure 3.4, we note that the BDE signature is able to describe partially occluded surfaces. A point on the tip of the right hand of the *Human* model is considered for an experiment where the BDE signature is generated, first when the whole body is intact and second, when the arm is separated from the rest of the body. The distances are normalized to ensure a proper comparison between the two BDE signatures. It is interesting to note that even when there is partial occlusion, the results are very similar until the scale parameter r reaches the boundary of the region where the arm was originally joined to the body.

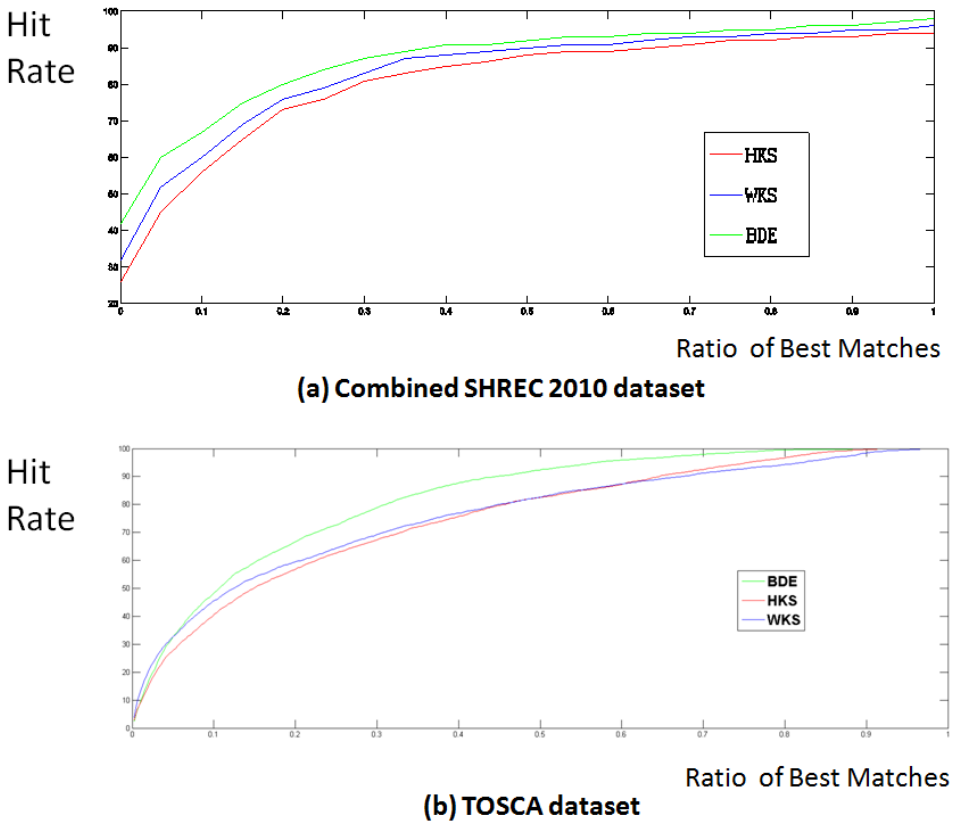

Figure 3.5: Hit rate graph showing the percentage of correct correspondences found in a subset of the k best matches (up to 1% of the total number of points) on (a) the entire SHREC 2010 benchmark dataset and, (b) the TOSCA benchmark data set which focuses primarily on isometric deformation, using the HKS, WKS and BDE signature.

3.5.3 Quantitative Analysis of Robustness

A comprehensive comparison of the proposed BDE descriptor to all surface feature descriptors in the literature is beyond the scope of this chapter. However, in the interest of making a quantitative assessment of the precision and robustness of the BDE descriptor, we compared the BDE to both, the WKS and HKS, on a shape matching problem. We performed our evaluation mainly on the SHREC 2010 benchmark dataset.

The feature points on the 3D shape surfaces are detected using the farthest point sampling

scheme [16] in feature space, i.e., sequential detection of the most distinguished feature points. Note that the basic idea underlying this feature detection technique is similar to that described in [16] except that it is performed in feature space in a manner similar to that described in [AUBRY11]. The feature points typically coincide with the characteristic surface locations which are chosen as reference points. For each of these reference points, the k best matches on the deformed shape are selected, where k varies from 1 to $N/100$ and N is the number of vertices on the deformed mesh. The hit rate, i.e., the percentage of feature points for which the correct corresponding point is among the k best matches, is plotted (on the y-axis) as a function of the percentual scale $100 \times k/N$ (on the x-axis) which allows for comparisons between meshes of different resolutions.

Experimental evaluation on the TOSCA [6] and SHREC 2010 [7] benchmark datasets allows us to quantify the robustness of the proposed BDE descriptor using the hit rate as a quality measure. The graph in Figure 3.5(a), shows the performance results on the entire SHREC 2010 benchmark dataset which encompasses nine different classes of transformations, i.e., isometry, topology, small and big holes, global and local scaling, noise, shot noise, and sampling. The graph in Figure 3.5(b), on the other hand, shows the performance results on the TOSCA [6] benchmark dataset which focuses primarily on cases exhibiting isometric deformations. In both cases, the BDE signature is seen to exhibit a superior hit rate compared to the HKS and WKS. Likewise, the graph in Figure 3.6 summarizes the performance of the BDE signature for four different classes of deformation on the SHREC 2010 benchmark dataset. The BDE signature displays a high hit rate for all four classes of deformation. The hit ratio can be seen to be $\geq 80\%$ for when the ratio of best matches ≥ 0.2 over all four deformation classes. The results depicted in Figure 3.5(a), (b) and Figure 3.6 are based on computation of the BDE signature using simple averaging.

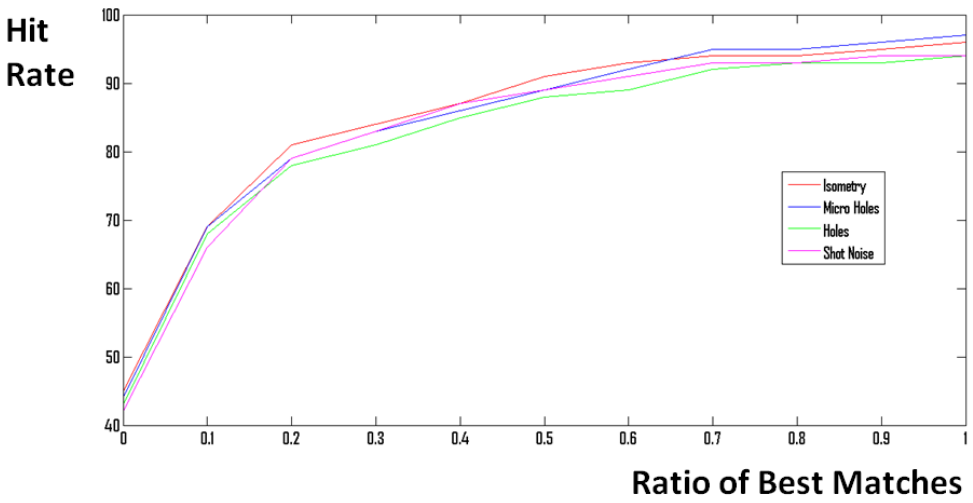

Figure 3.6: Hit rate graph showing the percentage of correct correspondences found in a subset of the k best matches (up to 1% of the total number of points) using the BDE signature on the SHREC 2010 benchmark dataset under four different classes of deformation.

3.5.4 Quantitative Comparison of Averaging Schemes

A quantitative comparison of the two averaging schemes, i.e., simple averaging and Gaussian averaging, used in the computation of the BDE descriptor was performed on the TOSCA benchmark dataset in a manner similar to that described in [AUBRY11]. The quantitative measurement procedure performed here is the same as the one described in Section 3.5.3. The graph in Figure 3.7 compares the performance of the two averaging schemes on the TOSCA [6] benchmark dataset. In this case, simple averaging is seen to exhibit a superior hit rate compared to Gaussian averaging. Consequently, in the remainder of the chapter, all experiments are performed using the simple averaging technique for BDE computation.

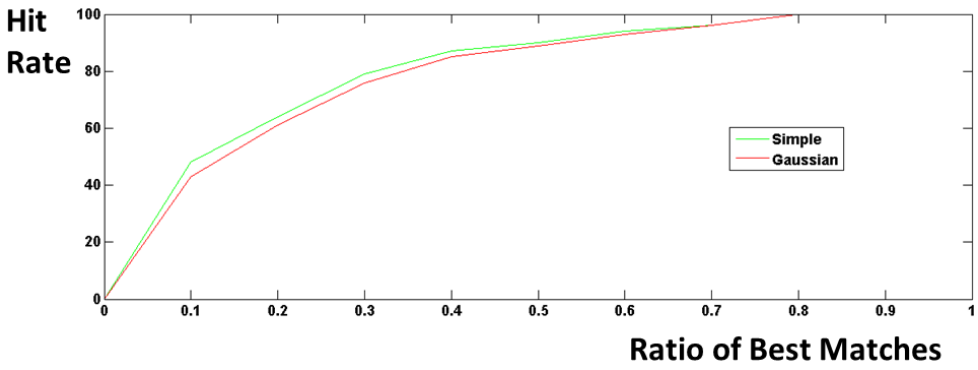

Figure 3.7: Hit rate graph showing the percentage of correct correspondences found in a subset of the k best matches (up to 1% of the total number of points) on the TOSCA benchmark data set using the simple and Gaussian averaging techniques for BDE computation. The TOSCA benchmark data set focuses primarily on isometric deformation.

3.5.5 Performance Statistics

All of the experiments reported in this chapter were performed on an Intel Core™ 3.4 GHz machine with 24 GB RAM. Table 5.4 reports the timing results for the proposed BDE signature.

Table 3.1: Timing results (in seconds) for meshes of different sizes.

Number of Points	Average Time (secs)
3000	9
24000	22
50000	39

3.5.6 Performance in Presence of Noise

It is critically important for a newly proposed shape signature to perform well in the presence of varying levels of noise. We performed a quantitative evaluation of the proposed BDE signature on the TOSCA dataset [6] under varying levels of additive white Gaussian noise.

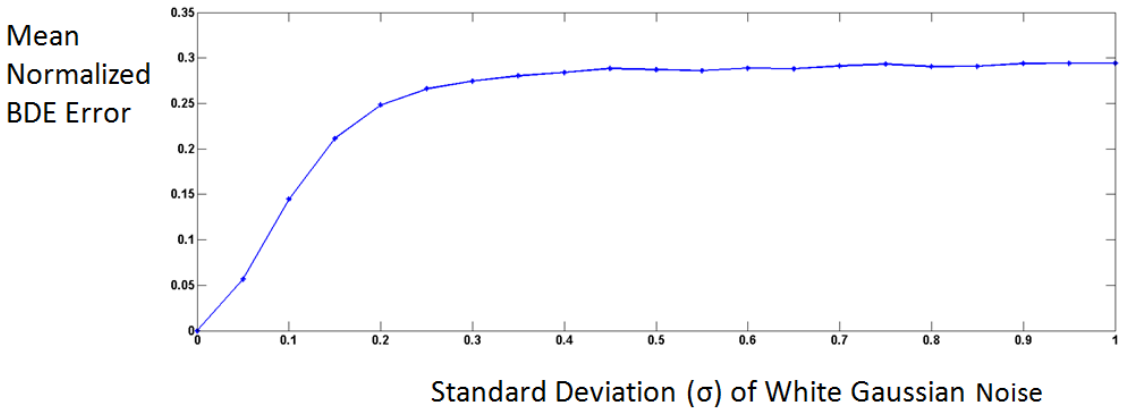

Figure 3.8: Robustness of the proposed BDE signature for varying levels of additive white Gaussian noise.

Each vertex on model surface is randomly displaced along the local surface normal. The displacement at each vertex is modeled by an independent and identically distributed (i.i.d.) random variable drawn from a zero-mean Gaussian distribution with a given value of σ (standard deviation). For each shape the feature points are chosen initially in the same manner as described in Section 3.5.3. For each feature point, the L^2 norm between the values of the original BDE signature and the noisy BDE signature is computed and normalized with respect to the original BDE signature value. The mean normalized error in the BDE signature values is computed over all the feature points for varying σ values, in the range $[0, 1]$, of the white Gaussian noise as depicted in Figure 4.9. It is interesting to note that the mean normalized error initially increases with increasing values of σ before asymptotically approaching a value of ≈ 0.3 for values of $\sigma > 0.3$, thus depicting the robustness of proposed BDE signature to additive white Gaussian noise.

3.6 Applications

3.6.1 Key Components Detection

The BDE signature can be used to identify key components on a deformable 3D surface. A key component is defined as a region containing several local discriminative features. Key components are useful in the sense that they are repeatable and robust against several transformations. Note that key components detection differs from mesh segmentation in that the former computes a *non-complete* decomposition of a mesh based on local component features [SIPIRAN12] whereas the latter aims at an *exhaustive* mesh partitioning. The idea behind key components detection is to decompose the mesh into a set of key components, which should be consistently detectable regardless of the transformation applied on the mesh.

Due to its high local discriminative ability, the key components are extracted at local scales. In the case of the BDE descriptor, the key components can be determined robustly using the BDE values at a local scale. In particular, we employ a k-means clustering based approach with $k = 2$ at a smaller scale (i.e., normalized scale parameter $r = 0.1$) in this set of experiments to cluster the surface points based on locally stable properties. The choice of $k = 2$ for the k-means clustering procedure ensures that the result is similar to that of an indicator function separating key components from non-key components whereas values of $k > 2$ would result in distinct key components some of which are symmetric under isometric deformation. The k-means clustering procedure results in a set of candidate vertices, which are locally discriminative and consistent across shapes which have undergone different isometric deformations as shown in Figure 3.9. Most importantly, these results suggest the repeatability of the key components detection technique even when the shapes have undergone isometric as well as non-isometric deformation, e.g. *Male* versus *Female* models shown in Figure 3.9.

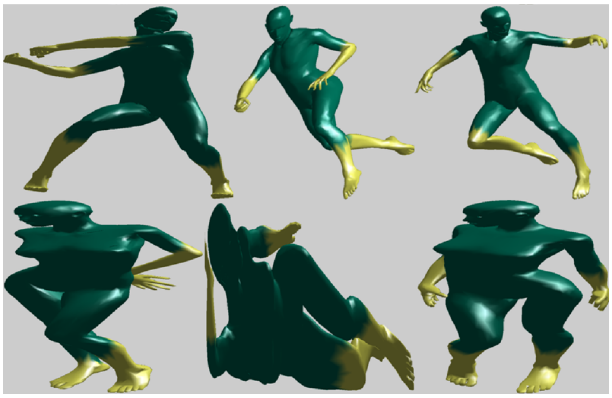

Figure 3.9: Key components detection (marked in yellow) using the BDE signature for two *Human* models *Male* (bottom row) and *Female* (top row) for different isometric deformations.

3.6.2 Sparse Corrrespondence Determination

A detailed treatment of the determination of complete or dense point-to-point correspondence between two deformable 3D shapes is beyond the scope of this chapter. However, we propose a means to determine the sparse point-to-point correspondence between two deformable 3D shapes using a recently developed spectral analysis technique [5]. The proposed sparse correspondence determination technique is based on the spectral analysis of each of the two deformable 3D shapes represented by a weighted shape graph $G(V, E, W)$. The vertices V of the graph G denote the point geometric features of the shape. The edge set E denotes the pair-wise relationships between each pair of vertices in the set V, making G a complete graph. The weight $w_{ij} \geq 0$ associated with an edge $(v_i, v_j) \in E$ encodes the affinity between the corresponding features represented by vertices v_i and v_j. The edge weights are represented by an $n \times n$ affinity matrix $W = [w_{ij}]_{i,j=1,2,\ldots,n}$.

The above formulation is extended for a joint graph as follows: Let $G_1(V_1, E_1, W_1)$ and $G_2(V_2, E_2, W_2)$ be the weighted graphs for shapes S_1 and S_2, respectively. The joint shape graph $G(V, E, W)$ is defined such that $V = V_1 \cup V_2$ and $E = E_1 \cup E_2 \cup (V_1 \times V_2)$ where $V_1 \times V_2$

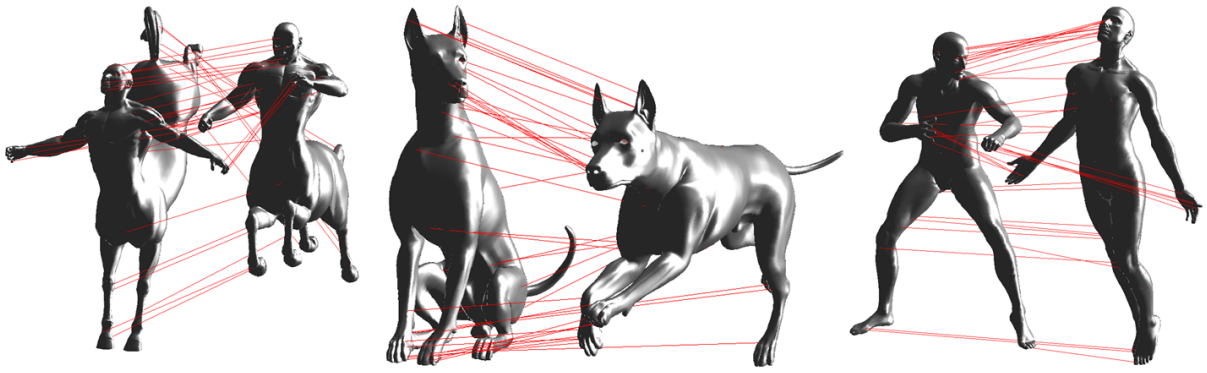

Figure 3.10: Results of sparse point correspondence determination via joint spectral embedding for three 3D deformable shape models: *Centaur, Dog* and *Human.*

is the set of edges connecting every pair of vertices in (V_1, V_2). The resulting joint affinity matrix W is given by:

$$W = \begin{bmatrix} W_1 & C \\ C^T & W_2 \end{bmatrix}_{(n_1+n_2)\times(n_1+n_2)} \tag{3.5}$$

The affinity submatrices W_1, W_2 and C are defined as follows:

$$(W_i)_{x,y} = \exp(-(\|f_i(x) - f_i(y)\|)^2) \tag{3.6}$$

$$C_{x,y} = \exp(-(\|f_i(x) - f_j(y)\|)^2) \tag{3.7}$$

where, $f_i(x)$ and $f_j(y)$ are the BDE signatures at points x and y on the two 3D shapes computed across multiple values of the scale parameter r and $\|\cdot\|$ denotes the Euclidean or L^2 norm. The joint spectral embedding of the BDE signatures ensures that geometrically similar surface points on two 3D shapes will have similar feature vectors in the jointly embedded space where the components of feature vectors represent the coefficients of the eigenvectors

of the joint affinity matrix W. Computation of the simple L^2 norm or Euclidean distance between the feature vectors in the jointly embedded space can generate good point-to-point correspondences across shapes as depicted in Figure 3.10.

3.7 Conclusions and Future Directions

A novel intrinsic geometric scale space signature, termed the Biharmonic Density Estimate (BDE) signature, has been proposed for isometry-invariant shape description. The BDE is shown to provide a formal means to study the geometric variability of shapes and to have potentially significant applications in computer graphics and computer vision. The applications of the BDE in the detection of key components on a deformable 3D surface and determination of sparse point correspondences between two deformable 3D shapes has also been demonstrated. In both of the aforementioned applications, it should be noted that we do not perform a semantic check of whether a computed match is one of the symmetric solutions, therefore the solution set obtained using the BDE descriptor will also contain symmetrically equivalent correspondence matches. In future, we plan to explore applications of the proposed multiscale BDE signature for large-scale content-based search and retrieval in 3D shape databases, point registration as well as for structural analysis, including partial and intrinsic symmetry detection and symmetrization.

Bibliography

[1] Aubry, M., Schlickewei, U., & Cremers, D. The wave kernel signature: a quantum mechanical approach to shape analysis, *Proc. IEEE Conf. Comp. Vision Pattern Recog.* (CVPR), 2011.

[2] Bansal M. & Daniilidis K., Joint spectral correspondence for disparate image matching, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 2802-2809, 2013.

[3] Belkin, M., Sun, J., & Wang, Y., Discrete Laplace operator on meshed surface, *Proc. Symp. Comp. Geometry*, pp. 278287, 2008.

[4] Belongie, S., Malik, J., & Puzicha, J. Shape context: A new descriptor for shape matching and object recognition, *Proc. Neural Info. Proc. Sys.*, 2000.

[5] Bronstein A.M., Bronstein M.M. & Kimmel R. Calculus of non-rigid surfaces for geometry and texture manipulation, *IEEE Trans. Visualization and Computer Graphics*, Vol. 13(5), pp. 902-913, 2007.

[6] Bronstein A.M., Bronstein M.M. & Kimmel R. *Numerical geometry of non-rigid shapes*, Springer, 2008.

[7] Bronstein A.M. *et al.*, SHREC 2010: robust large-scale shape retrieval benchmark, *Proc. Eurographics Workshop on 3D Object Retrieval* (3DOR), 2010.

[8] Bronstein M.M. & Bronstein A.M., Shape recognition with spectral distances, *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5), pp. 1065-1071, May 2011.

[9] Johnson A. & Hebert M., Using spin images for efficient object recognition in cluttered 3D scenes, *IEEE Trans. Pattern Anal. Mach. Intel.*, 21(5), pp. 433-449, 2002.

[10] Karni Z. & Gotsman C., Spectral compression of mesh geometry, *Proc. ACM SIGGRAPH*, 2000.

[11] Levy B., Laplace-Beltrami eigenfunctions: towards an algorithm that understands geometry, *Proc. IEEE Int. Conf. on Shape Model. Applic.*, pp. 13, 2006.

[12] Li X. & Guskov I., Multi-scale features for approximate alignment of point-based surfaces, *Proc. Eurographics Symp. Geom. Proc.*, pp. 217, 2005.

[13] Ling, H., & Okada, K. Diffusion distance for histogram comparison, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, (CVPR), Vol. 1, 2006.

[14] Lipman, Y., Rustamov, R. M., & Funkhouser, T. A. Biharmonic distance, *ACM Trans. Graphics*, 29(3), pp. 27:1-27:11, 2010.

[15] Manay, S., Cremers, D., Hong, B. W., Yezzi, A. J., & Soatto, S. Integral invariants for shape matching, *IEEE Trans. Pattern Anal. Mach. Intel.*, 28(10), pp. 1602-1618, 2006.

[16] Moenning C. & Dodgson N.A., Fast Marching farthest point sampling, *Proc. EUROGRAPHICS*, 2003.

[17] New, A. T., Mukhopadhyay, A., Arabnia, H. R., & Bhandarkar, S. M. Non-rigid shape correspondence and description using geodesic field estimate distribution, *Proc. ACM SIGGRAPH*, August 2012.

[18] Osada, R., Funkhouser, T., Chazelle, B., & Dobkin, D. Shape distributions, *ACM Trans. Graphics*, 21(4), pp. 807-832, 2002.

[19] Ovsjanikov, M., Mrigot, Q., Mmoli, F., & Guibas, L. One point isometric matching with the heat kernel, *Proc. Symp. Geometry Processing*, 2010.

[20] Peyre G., *Toolbox Graph*, July 2009.

[21] Pinkall U. & Polthier K., Computing discrete minimal surfaces and their conjugates, *Experim. Maths.*, 2(1), pp. 15-36, 1993.

[22] Reuter M., Wolter F.-E. & Peinecke N., Laplace-Beltrami spectra as "Shape-DNA" of surfaces and solids. *Computer-Aided Design*, 38(4), pp. 342-366, 2006.

[23] Rustamov R., Laplace-Beltrami eigenfunctions for deformation invariant shape representation, *Proc. Eurographics Symp. Geom. Proc.*, pp. 225-233, 2007.

[24] Sipiran I. & Bustos B., Key-component detection on 3D meshes using local features. *Proc. 5th Eurographics Conf. 3D Object Retrieval*, 2012.

[25] Sun, J., Ovsjanikov, M., & Guibas, L. A concise and provably informative multi-scale signature based on heat diffusion, *Comp. Graphics Forum*, 28(5), pp. 1383-1392, 2009.

[26] Taubin G., A signal processing approach to fair surface design, *Proc. ACM SIGGRAPH*, 1995.

[27] Wardetzkey M., Convergence of the cotangent formula: An overview. *Discrete Differn. Geom.*, Birkhuser Basel, 2005, pp. 89-112.

[28] Xu G., Discrete Laplace-Beltrami operators and their convergence, *Comput. Aided Geom. Des.*, 21(8), pp. 767-784, 2004.

[29] Yen, L., Fouss, F., Decaestecker, C., Francq, P., & Saerens, M. Graph nodes clustering based on the commute-time kernel, *Proc. 11th Pacific-Asia Conf. Knowledge Discovery and Data Mining* (PAKDD), 2007.

[30] Zou, G., Hua, J., Lai, Z., Gu, X., & Dong, M. Intrinsic geometric scale space by shape diffusion, *IEEE Trans. Visual. and Comp. Graphics*, 15(6), pp. 1193-1200, 2009

Chapter 4

Intrinsic Symmetry Detection and Characterization

4.1 Introduction

Symmetry is ubiquitous in nature and also in artificial man-made objects. The detection and characterization of shape symmetry has attracted much attention in recent times, especially within the computer graphics community [MITRA13]. Although most of the existing literature has focused on the detection of *extrinsic* symmetries, a popular approach being transformation space voting [MITRA06], there has been steadily growing interest in detection and characterization of *intrinsic* symmetries. Most recent efforts in intrinsic symmetry detection have focused on detection of global symmetries [LIPMAN10a]; [OVSJANIKOV08].

Informally, symmetry can be considered as a *distance preserving* transformation. For a symmetry transformation T, the distance between two points p and q on the surface is equal to the distance between transformed points $T(p)$ and $T(q)$ on the surface. In the case of extrinsic symmetry, the distance measure used is *Euclidean*, whereas in the case of intrinsic symmetry, the distance measure used is *intrinsic* in nature such as, the *geodesic distance*

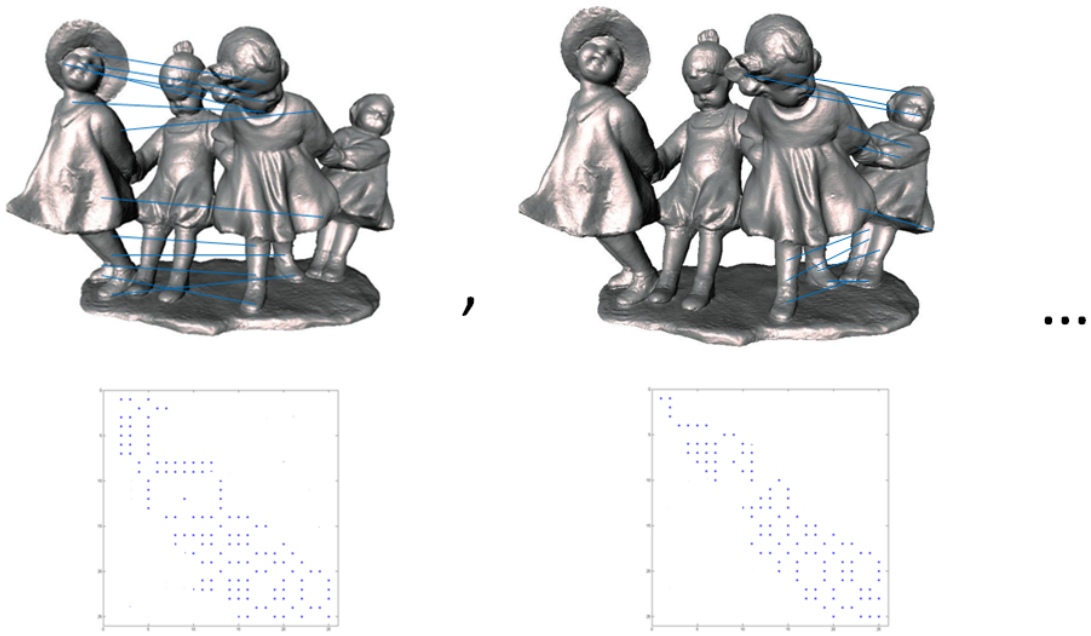

Figure 4.1: Symmetry extraction in functional space. The top row depicts two significant symmetry transformations along with a few representative point correspondences. The corresponding functional map matrices are shown in the bottom row.

measure or the *biharmonic distance measure*. It is generally recognized that detection of overlapping intrinsic symmetry is a more challenging problem due to the significantly larger search space involved in the detection of symmetric regions (when compared to global symmetry analysis) and in the determination of symmetry revealing transforms (when compared to extrinsic symmetry detection). The importance of the overlapping intrinsic symmetry detection and characterization problem is further underscored by the fact that it is more general in nature, since the extrinsic symmetry detection problem can be considered a special case of overlapping intrinsic symmetry detection.

We present a formal definition of overlapping intrinsic symmetry. An intrinsic symmetry over a shape is a subregion with associated self-homeomorphisms that preserve all pairwise

intrinsic distances [MITRA13]. Complex shapes often exhibit multiple symmetries that overlap spatially and vary in visual appearance as depicted in Figure 4.1. Overlapping symmetry analysis enables the construction of high-level representations that enhance the understanding of the underlying shape and facilitate solutions to such problems as shape correspondence, shape editing, and shape synthesis [WANG11]. However, analysis of overlapping symmetry poses additional challenges as described below.

Existing approaches to intrinsic symmetry detection, including those based on region growing [XU09], partial matching [RAVIV10], and symmetry correspondence [LIPMAN10a] are unable to extract physically overlapping symmetries. Lipman et al. [LIPMAN10a] cluster sample surface points from an input shape and use a symmetry correspondence matrix (SCM) to identify intrinsic symmetry properties of groups of surface points. In their approach, each SCM entry measures the extent of symmetry between two surface points based on some measure of intrinsic geometric similarity between the local neighborhoods of the points. Xu et al. [XU12] let surface point pairs vote for their partial intrinsic symmetry and perform intrinsic symmetry grouping using a 2-stage spectral clustering procedure. However, their approach lacks the ability to retrieve the final symmetry map which makes characterization of the specific intrinsic symmetry a difficult problem.

The key idea behind the proposed scheme for intrinsic symmetry detection and characterization is to approach the problem from a shape correspondence perspective and generate the transformation map which can further be used to describe the symmetry space. To this end, we perform two stages of processing where in the first stage, representative symmetric point pairs are identified based on their local geometry and a global distance representation and in the second stage the original transformation is retrieved as a *map* to facilitate further characterization of the underlying symmetry. The detected intrinsic symmetries are quite representative and the corresponding functional matrices exhibit a high degree of diagnonality, as shown in Figure 4.1.

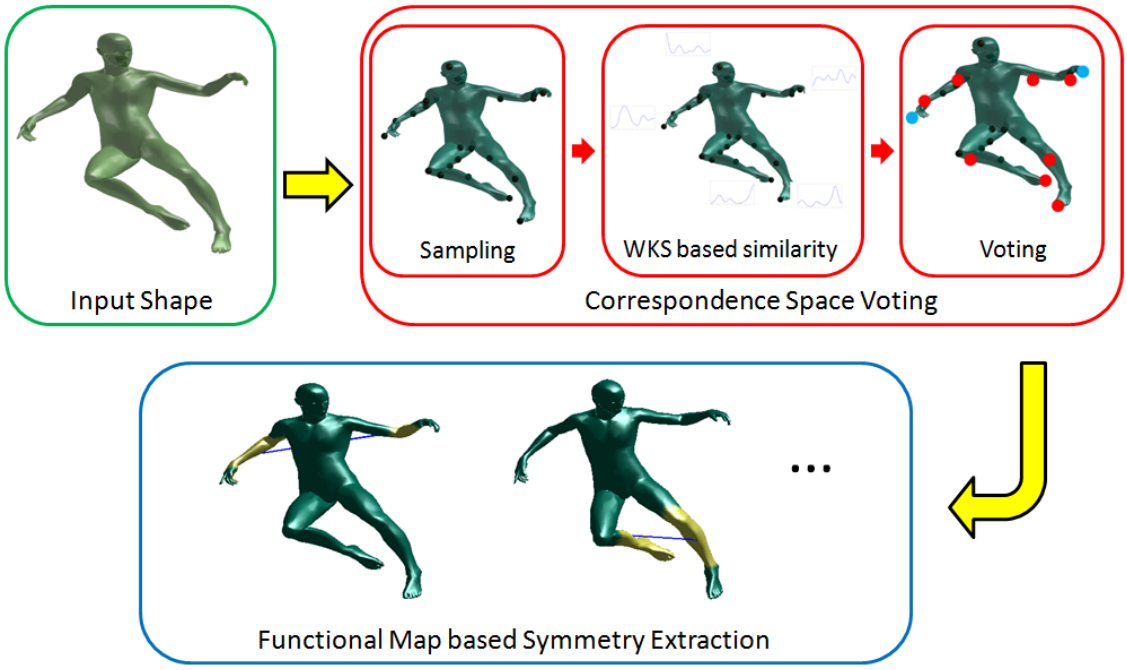

Figure 4.2: Overview of the proposed symmetry detection and characterization framework.

4.1.1 Overview

The input to the proposed symmetry detection and characterization algorithm is a 3D shape that is approximated by a 2-manifold triangular mesh. The local intrinsic geometry is quantified using the *Wave Kernel Signature* (WKS) [AUBRY11] whereas the global intrinsic geometric distance measure chosen is the *Biharmonic Distance Measure* (BDM) [LIPMAN10b]. The proposed algorithm consists of a voting procedure performed in correspondence space defined over a set of locally symmetric surface point pairs sampled from the input 3D shape, followed by the generation of a functional map as depicted in Figure 4.2.

Correspondence Space Voting A fundamental question in symmetry detection is quantifying the extent of symmetry between a pair of points. The primary challenge in identifying

potentially symmetric point pairs is to come up with conditions strong enough to adequately constrain the symmetry search space such that the symmetry detection procedure is computationally tractable. We rely on a local criterion, such as geometric similarity, and a global criterion, such as the distance-based symmetry support received by a point pair $\{a, b\}$, to detect and quantify the extent of symmetry between points a and b. We refer to a point pair $\{a, b\}$ which satisfies the local geometric similarity criterion as a *good voter*. The point pairs within the population of *good voters*, that enjoy sufficiently strong global distance-based symmetry support are deemed to be *symmetric* point pairs. The global symmetry support for a point pair $\{a, b\}$ is quantified by the number of other point pairs which potentially share the intrinsic symmetry properties of $\{a, b\}$. The global symmetry support is computed using a simple, distance-based symmetry criterion defined over two point pairs within the population of *good voters* as shown in Figure 4.2.

The input to the correspondence space voting procedure comprises of point pairs that can be considered as candidates for symmetry. In the initial step of correspondence space voting procedure, we sample a set of locally symmetric surface point pairs from the input 3D shape based on the similarity of their WKS values. To estimate the symmetry support received by a point pair, we perform a voting procedure by counting the number of *good voters* which potentially share the same intrinsic symmetry as the source point pair. The voting procedure ensures that we have a set of good point pair initializations from which we can create an initial map of the symmetry transformation that can then be extrapolated to other surface point pairs on the 3D shape.

Computation of the Functional Map A functional map provides an elegant representation of the maps between surfaces, allowing for efficient inference and manipulation. In the functional map approach, the concept of a map is generalized to incorporate correspondences between real valued functions rather than simply between surface points on the 3D

shapes. Our choice of the multi-scale eigen-basis of the Laplace-Beltrami operator makes the functional map representation both, very compact and yet informative. We show that the proposed functional map formulation not only results in a compact description of the underlying map of the symmetry transformation, but also enables meaningful characterization of the symmetry transformation.

The primary contributions of our chapter can be summarized as follows:

1. Exploiting the functional map representation in conjunction with the correspondence space voting technique of Xu et al. [XU12]. The voting technique in Xu et al. [XU12] is followed by a computationally complex optimization step comprising of a two-stage clustering procedure for symmetry detection. The proposed technique, on the other hand, is based on detection of a few initial symmetric candidates followed by a computationally efficient generalization procedure for overall symmetry detection based on the functional map representation and solving a system of linear equations.

2. Robust and meaningful characterization of the symmetry transformation via formulation of a symmetry space which allows one to *quantitatively* distinguish between instances of simple and complex intrinsic symmetry. To the best of our knowledge, this quantitiative distinction has not been attempted in the published literature.

3. Enabling recovery of the symmetry groups via clustering on the functional maps.

4.2 Related Work

The research literature on symmetry detection has grown substantially in recent years as shown in the recent comprehensive survey by Mitra et. al. [MITRA13]. In this chapter, we do not attempt to provide an *exhaustive* exposition of the state of the art in symmetry detection; rather we restrict our discussion to existing works that are most closely related to

the proposed approach.

4.2.1 Symmetry detection in transformation space

Several recent approaches to detect approximate and partial extrinsic symmetries have focused on algorithms that are based on clustering of *votes* for symmetry detection in a parameterized *transformation space* [IMIYA99], [MITRA06], [YIP00], [27]]. For example, Mitra et al. [MITRA06] generate votes in a transformation space to align pairs of similar points and then cluster these votes in a space spanned by the parameters of the potential symmetry transformations. Regardless of how good the shape descriptors are, the aforementioned methods are not effective at finding correspondences between points in complex symmetry orbits that are spread across multiple distinct clusters in the transformation space. Since the dimensionality of the transformation space increases with the complexity of the symmetry, the voting procedure in transformation space becomes increasing intractable when dealing with complex symmetries.

4.2.2 Symmetry representation in transformation space

There exists a substantial body of published research literature that characterizes shape representations based on the extent of symmetry displayed by an object with respect to multiple transformations. Kazdhan et al. [KAZDHAN03] have proposed an extension of Zabrodsky's symmetry distance [ZABRODSKY95] in their formulation of characteristic functions, resulting in a *symmetry descriptor* that measures the symmetries of an object with respect to all planes and rotations through its center of mass. Podolak et al. [PODOLAK06] have extended the symmetry descriptor to define a *planar reflective symmetry transform* (PRST) that measures reflectional symmetries with respect to all planes in 3D space. Rustamov et al. [RUSTAMOV08] have extended the PRST by incorporating more information in the

symmetry transformations to obtain a family of symmetry transforms indexed by a single parameter, namely the surface point-pair correlation values at multiple radii. Although the above representations provide a measure of symmetry for a regularly sampled set of transformations within a group, they are practical only for transformation groups of low dimensionality (for example, rigid body transformations would require one to store a six-dimensional matrix) and break down when faced with groups of higher dimensionality.

4.2.3 Discovery of repeating structures

There exists a class of techniques that exploits the redundancy in repeating structures to robustly detect symmetries [BOKELOH09]; [LI06]; [LIU07]; [PAULY08]; [SHIKHARE01]. The transformation space voting method proposed by Mitra et al. [MITRA06] is extended in [PAULY08] by fitting parameters of a transformation generator to optimally register the clusters of votes in transformation space. Berner et al. [BERNER08] and Bokeloh et al. [BOKELOH09] have taken a similar approach using subgraph matching of feature points and feature lines, respectively, to establish potential correspondences between repeated structures. The subgraph matching procedure is followed by an *Iterative Closest Point* (ICP) algorithm to simultaneously grow the corresponding regions and refine the matches over all detected patterns. The schemes proposed by Berner et al. [BERNER08] and Bokeloh et al. [BOKELOH09] allow the detection of repeated patterns even in noisy data [PAULY08], but at the cost of requiring a-priori knowledge of the commutative group expected in the data. Also, the non-linear local optimization procedure within the ICP algorithm could cause it to get trapped in a local minimum if the initialization is not good enough.

4.2.4 Eigen-analysis methods

Lipman et al. [LIPMAN10a] have proposed an eigen-analysis technique for symmetry detection that relies on spectral clustering. The eigenvectors, corresponding to the largest eigenvalues of their geometric similarity-based symmetry correspondence matrix (SCM), characterize the symmetry-defining orbits, where each orbit includes all points that are symmetric with one another. However, their work is not well suited for multi-scale partial symmetry detection. First, expressing local point similarities as symmetry invariants is only appropriate for global intrinsic symmetry detection. In the case of partial symmetry detection, it is not always possible to reliably judge if two surface points are symmetric by comparing only their point (i.e., local) signatures, especially when one point lies on the boundary of symmetric regions. Moreover, their single-stage clustering procedure is unable to identify overlapping symmetries.

Xu et al. [XU12] have extended the eigen-analysis approach of Lipman et al. [LIPMAN10a] by incorporating the concept of global intrinsic distance-based symmetry support accompanied by a 2-stage spectral clustering procedure to distinguish between scale determination and symmetry detection. Although they showed some interesting results, the 2-stage spectral clustering procedure caused their method to be extremely slow. Furthermore, the absence of transformation map retrieval meant that further processing of the detected symmetries, which are represented as point pairs, was extremely inefficient.

The proposed scheme is decoupled into two stages termed as *Correspondence Space Voting* (CSV) and *Transformation Space Mapping* (TSM). Although the CSV procedure is inspired by the work of Xu et al. [XU12], we have bypassed the two computationally intensive procedures of spectral clustering and all-pairs geodesic distance computation in their scheme [XU12] to improve significantly the running time of the proposed CSV procedure. Moreover, our incorporation of the TSM procedure in symmetry detection is, to the best of our knowledge, novel in that it not only provides a concise *description* of the underlying sym-

metry transformation, but also enables formal *characterization* of the underlying symmetry transformation.

4.3 Theoretical Framework

In this section, we present a formal description of the problem being addressed in this chapter. In particular, we define the input to and output of the proposed algorithm. We also define the type of intrinsic symmetries the proposed algorithm is designed to detect and the formal characterization of these symmetries. There are two primary aspects to the theoretical framework for the proposed algorithm, i.e., *Correspondence Space Voting* (CSV) and *Transformation Space Mapping* (TSM) or *Functional Map Retrieval* (FMR). The terms TSM and FMR are used interchangeably in this chapter. In the case of CSV, a joint criterion, that combines local intrinsic surface geometry and global intrinsic distance-based symmetry, is proposed and shown to result in a provably necessary condition for intrinsic symmetry. In the case of TSM or FMR, a formal scheme is proposed for the characterization of the complexity of the detected symmetries, and its effectiveness demonstrated.

We have restricted our study of intrinsic symmetry primarily to isometric involutions for two reasons; first, the restriction results in a provably necessary condition for intrinsic symmetry which provides theoretical soundness and second, the proposed symmetry criterion bounds the symmetry search space sufficiently, ensuring that the solution is computationally tractable [XU12]. But, unlike [XU12], the computationally complex CSV procedure is only an initialization step in the proposed method. Following the initial detection of good symmetry correspondences using the CSV procedure, we proceed to retrieve the map of the symmetry transformation by leveraging the functional map framework. The generalization of the symmetry detection over the entire surface is performed by solving a system of linear equations which not only returns the underlying map of the symmetry transformation, but

also ensures the scalability of our method with respect to the number of surface points. The intrinsic symmetry criterion ensures that the retrieved functional matrices exhibit a diagonality characteristic that can be quantified using a cost matrix specifically designed for this purpose. The inner product of the functional matrix with the cost matrix is shown to result in a *quantitative* measure of the complexity of the detected symmetry.

Before describing the proposed algorithm in detail, we summarize the previously developed concepts, tools and techniques that are used in the proposed algorithm.

4.3.1 Biharmonic Distance Measure

Measuring intrinsic distances between pairs of points on a 3D surface is a fundamental problem in computer graphics. The most important properties of an intrinsic distance are that it is a metric, smooth, locally isotropic, globally shape-aware, isometry invariant, insensitive to noise and small topology changes, parameter-free, and practical to compute on a discrete mesh. But most existing intrinsic distance measures do not follow the above properties. The *Biharmonic Distance Measure* (BDM) is a welcome exception amongst the intrinsic distance measures encountered in the published literature since it satisfies all the desirable properties mentioned above [LIPMAN10a].

The BDM is similar in form to the *Diffusion Distance Measure* (DDM) [LING06]. The BDM kernel is based on the Green's function of the biharmonic differential equation. In the continuous case, the (squared) biharmonic distance between two points x and y can be defined using the eigenvectors (ϕ_k) and eigenvalues (λ_k) of the Laplace- Beltrami operator [LIPMAN10a] as follows:

$$d_M(x,y)^2 = \sum_{k=1}^{\infty} \frac{(\phi_k(x) - \phi_k(y))^2}{\lambda_k^2} \tag{4.1}$$

The above definition of the BDM is slightly different from that of the DDM where the

denominator in the case of the DDM is $e^{2t\lambda_k}$. However, this subtle change ensures greater control over the characterization of the global and local properties of the underlying manifold in the case of the BDM. Consequently, the BDM is fundamentally different from the DDM with significantly different properties.

The BDM, as expressed in equation (4.1), captures the rate of decay of the normalized eigenvalues λ_k of the Laplace- Beltrami operator; if the decay is too slow, it produces a logarithmic singularity along the diagonal of the Green's function [YEN07]. Alternatively, too fast a decay basically ignores the eigenvectors associated with higher frequencies, resulting in the BDM being global in nature (i.e., the local surface details are ignored). Lipman *et al.* [LIPMAN10b], demonstrated that performing quadratic normalization provides a good balance, ensuring that the decay is slow enough to capture the local surface properties around the source point and yet rapid enough to encapsulate global shape information.

In particular, Lipman *et al.* [LIPMAN10b] have theoretically proved two important properties of the BDM, i.e., that it is (i) a metric, and (ii) smooth everywhere except at the source point where it is continuous. The key observation is that for 3D surfaces, the eigenvalues λ_k, $k = 1, 2, \ldots$, of the Laplacian are an increasing function of k resulting in the continuity of the BDM everywhere and also the smoothness of the BDM everywhere except at the source point, where it has only a derivative discontinuity.

4.3.2 Wave Kernel Signature

The comprehensive analysis of shapes entails the formulation of a geometric feature descriptor which characterizes each point on the object surface taking into account its relation to the entire shape. The Wave Kernel Signature (WKS) of a surface point x, is a geometric feature descriptor which evaluates the probability of a quantum particle with a certain energy distribution to be located at point x. The behavior of the quantum particle on the surface is governed by the Schrodinger equation [SCHRODINGER26]. Assuming that the quantum

particle has an initial energy distributed around some nominal energy with a probability density function $f(e)$, the solution of the Schrodinger equation can then be expressed in the spectral domain as:

$$\psi(x,t) = \sum_{k>=1} e^{ie_k t} f(e_k)\phi_k(x) \tag{4.2}$$

Aubry et al. [AUBRY11] consider a family of log-normal energy distributions centered around some mean log energy $\log e$ with variance σ^2. This particular choice of distributions is motivated by a perturbation analysis of the Laplacian spectrum [AUBRY11]. Having fixed the family of energy distributions, each point x on the surface is associated with a WKS of the form: $p(x) = (p_{e_1}(x), \ldots, p_{e_n}(x))^T$ where $p_{e_i}(x)$ is the probability of measuring a quantum particle with the initial energy distribution $e_i(x)$ at point x. Aubry et al. [AUBRY11] use logarithmic sampling to generate the values $e_1(x), \ldots, e_n(x)$.

The WKS can be shown to exhibit a band-pass characteristic. This reduces the influence of low frequencies and allows better separation of frequency bands across the descriptor dimensions. As a result, the WKS descriptor exhibits superior feature localization compared to the *Heat Kernel Signature* (HKS) [LITMAN13].

4.3.3 Functional Map

A functional map [OVSJANIKOV12] is a novel approach for inference and manipulation of maps between shapes that tries to resolve the issues of correspondences in a fundamentally different manner. Rather than plotting the corresponding points on the shapes, the mappings between functions defined on the shapes are considered. This notion of correspondence generalizes the standard point-to-point map since every point-wise correspondence induces a mapping between function spaces, while the opposite, in general, is not true [OVSJANIKOV12].

The functional map framework described above provides an elegant means, using a func-

tional representation, for avoiding the direct representation of correspondences as mappings between shapes. Ovsjanikov et al. [OVSJANIKOV12] have noted that when two shapes X and Y are related by a bijective correspondence $t : X \to Y$, then for any real function $f : X \to \mathbb{R}$, one can construct a corresponding function $g : Y \to \mathbb{R}$ as $g : f \circ t^{-1}$. In other words, the correspondence t uniquely defines a mapping between the two function spaces, i.e., $F(X, \mathbb{R}) \to F(Y, \mathbb{R})$, where $F(X, \mathbb{R})$ denotes the space of real functions on X. Equipping X and Y with harmonic bases, $\{\phi_i\}_{i\geq 1}$ and $\{\psi_j\}_{j\geq 1}$, respectively, one can represent a function $f : X \to \mathbb{R}$ using the set of (generalized) Fourier coefficients $\{a_i\}_{i\geq 1}$ as $f = \sum_{i\geq 1} a_i\phi_i$. Translating this representation into the other harmonic basis $\{\psi_j\}_{j\geq 1}$, one obtains a simple representation of the correspondence between the shapes given by $T(f) = \sum_{i,j\geq 1} a_i c_{ij}\psi_j$ where c_{ij} are Fourier coefficients of the basis functions of X expressed in the basis of Y, defined as $T(\phi_i) = \sum_{i,j\geq 1} c_{ij}\psi_j$. The correspondence t between the shapes can thus be approximated using k basis functions and encoded using a $k \times k$ matrix $C = (c_{ij})$ of these Fourier coefficients, referred to as the functional matrix. In this representation, the computation of the shape correspondence $t : X \to Y$ is translated into a simpler task of determining the functional matrix C from a set of correspondence constraints. The matrix C has a diagonal structure if the harmonic bases $\{\phi_i\}_{i\geq 1}$ and $\{\psi_j\}_{j\geq 1}$ are *compatible*.

Having discussed the previously developed concepts, tools and techniques, we describe the proposed symmetry detection and characterization algorithm in detail.

4.3.4 Input to the Proposed Algorithm

The problem domain under consideration is a compact, connected 2-manifold, M, with or without a boundary. The manifold M is a *3D shape*, i.e., $M \to \mathbb{R}^3$. Distances on the manifold M are expressed in terms of an intrinsic distance measure. In particular, the BDM is used on account of its ease of calculation and greater robustness to local surface perturbations when compared to the geodesic distance measure [LIPMAN10b]. In the remainder of the

chapter, we use the term intrinsic distance and biharmonic distance interchangeably.

4.3.5 Output of the Proposed Algorithm

Since a symmetry transformation is represented by a functional map in the proposed algorithm, the output of the proposed algorithm consists of functional maps represented as matrices. The output can be regarded as a complete description of all the overlapping intrinsic symmetries represented in a compact and informative manner.

4.3.6 Definition of Intrinsic Symmetry

We assume that we are given a compact connected 2-manifold M without boundary. Following [RAVIV07] we deem manifold M to be intrinsically symmetric if there exists a homeomorphism $T\colon M \to M$ on the manifold that preserves all intrinsic distances. That is,

$$d_M(p,q) = d_M(T(p), T(q)) \forall p, q \in M \tag{4.3}$$

where $d_M(p,q)$ is the intrinsic distance between two points on the manifold. In this case, we call the mapping T an intrinsic symmetry.

4.3.7 Symmetry Criteria

We propose two simple criteria to test whether two surface point pairs $\{x, x'\}$ and $\{y, y'\}$ on the manifold M potentially share the same intrinsic symmetry. The first criterion, which is based on local intrinsic geometry, determines the symmetry potential of the two surface point pairs by comparing their corresponding WKS values as follows:

$$\begin{aligned} WKS(x,t) &= WKS(x',t) \text{ and} \\ WKS(y,t) &= WKS(y',t)\ \forall t \geq 0 \end{aligned} \tag{4.4}$$

where t is a scale parameter. Although local intrinsic geometric similarity is not a sufficient criterion, it is a necessary one [OVSJANIKOV10] and when imposed serially along with the next criterion, ensures a smaller search space. The second criterion is based on intrinsic distance as follows:

$$\begin{aligned} d_M(x,y) &= d_M(x',y') \text{ and} \\ d_M(x,y') &= d_M(y,x') \end{aligned} \tag{4.5}$$

This criterion is motivated by [XU12] and can be considered a theoretically sufficient condition. The above two criteria are necessarily satisfied if the surface point pairs under consideration correspond to the same intrinsic symmetry.

4.3.8 Correspondence Space Voting (CSV)

The first step in the proposed symmetry detection and characterization framework is the CSV procedure. Although a voting procedure has been previously incorporated in an earlier symmetry detection technique, it was carried out primarily in transformation space [MITRA06]. The importance of the correspondence space for the detection of symmetry was explained more recently by Lipman et al. [LIPMAN10a]. In our case, although the detected symmetry is finally represented in functional space, it is critical to have good initialization to ensure the success of the final map generation. We have designed and implemented a CSV algorithm to facilitate good initial guesses. But, unlike [XU12], the computationally complex global CSV procedure is performed in a reduced search space resulting from employing an efficient sampling strategy and the local geometric similarity criterion (equation (4.4)). The generalization of the symmetry detection over the entire shape is performed by solving a system of linear equations in a computationally efficient manner. The proposed CSV algorithm comprises of three stages described in the following subsections:

Figure 4.3: Sampling results on three different *human* shapes

Selection of Surface Points

A subset of points with adequate discriminative power needs to be selected prior to the generation of surface point pairs to ensure the reduced space for voting. A subset X consisting of n sample points is chosen from the surface of the given input 3D shape using the *farthest point random sampling* strategy [ELDAR97]. Although originally designed to operate on geodesic distances generated using the *Marching Cubes* algorithm [LORENSEN87], in our case we have employed the farthest point random sampling strategy in biharmonic distance space. The results of the sampling procedure are depicted in Figure 4.3.

In our implementation of the farthest point random sampling strategy, a single point is selected randomly at first and the remaining points are chosen iteratively from the remainder set by selecting the farthest point in the biharmonic distance space at each iteration. This strategy generates a set of points located mostly in the vicinity of the shape extrema which can then be used in the subsequent surface point pair generation procedure.

Generation of Surface Point Pairs

From the chosen subset X consisting of n sample points, the surface point pairs are generated by exploiting the similarity of their local intrinsic geometric structure. The similarity of local intrinsic geometric structure of two surface points is determined by comparing their corresponding WKS values.

Although the WKS is invariant under isometric deformation, in most practical cases where the underlying surface is represented by a discrete triangular mesh, it is not possible to satisfy strictly the invariance criterion in equation (4.4). Consequently, we have considered the 2-norm of the WKS (i.e., simply the squared Euclidean distance in the WKS space) and chosen surface point pairs from the subset X which satisfy the WKS invariance criterion in equation (4.4) to within a prespecified threshold T_{WKS} instead of requiring strict equality. The surface point pairs that satisfy the WKS invariance criterion in equation (4.4) are considered for inclusion in the next step of global distance-based voting. This relaxation of the WKS invariance criterion ensures that even cases of overlapping intrinsic symmetries are considered in the voting procedure. We have coined the term *good voters* to denote the subset of surface point pairs that satisfy the WKS invariance criterion in equation (4.4) to within the pre-specified threshold.

Global Distance-based Voting

The global distance-based voting step in the proposed symmetry detection technique is inspired by the work of Xu et. al. [XU12]. A subset of symmetric point pairs is extracted from the set of *good voters* using the global distance-based voting procedure prior to functional map generation. The goal of the voting procedure is to accumulate symmetry support for the *good voters* and extract the symmetric point pairs based on the level of symmetry support received.

A point pair in the set of *good voters* is deemed to be symmetric if it has a sufficiently large global symmetry support, which in [XU12] is measured by the number of point pairs that satisfy the intrinsic distance criterion in equation (4.5). Xu et al. [XU12] have presented a voting technique based on two point pairs followed by spectral clustering on the symmetric point pairs which enables one to distinguish whether a point pair supports one particular symmetry or more than one type of symmetry. However, since we are not interested, at this

stage, in making the above distinction between the point pairs, we have chosen to adopt a straightforward approach of choosing a point pair and letting the set of *good voters* vote and decide whether or not the particular point pair satisfies the intrinsic distance criterion in equation (4.5). Another modification to the previous voting procedure of Xu et al. [XU12] is our use of the biharmonic distance as the intrinsic distance instead of the geodesic distance. This modification provides two advantages over the previous voting procedure. First, since the computation of biharmonic distances between all surface point pairs is more efficient than the computation of geodesic distances, the proposed voting procedure is much faster. Second, biharmonic distances are less sensitive to noise and surface perturbations compared to geodesic distances, making the proposed voting procedure more robust.

4.3.9 Transformation Space Mapping and Symmetry Extraction Using Functional Maps

In the proposed symmetry extraction algorithm, instead of comparing two different shapes, we propose to compare two symmetric regions within the same shape. In particular, based on the previously detected set of symmetric point pairs, we leverage the functional map representation for symmetry extraction. For each pair of symmetric points, we deem one point as the source and the other as the destination and choose a local region around each point. The ordering of the source and destination points within the pair is the same as the one originally chosen during the voting procedure. The corresponding eigen-bases for the points in the source and destination regions are computed. These eigen-bases are ordered based on their similarity with each other and the final functional map for that particular symmetry is extracted. The functional map representation ensures that (a) the problem of symmetry extraction is tractable and, (b) the resulting symmetry can be represented, not by a large matrix of point correspondences, but rather as a more compact functional map

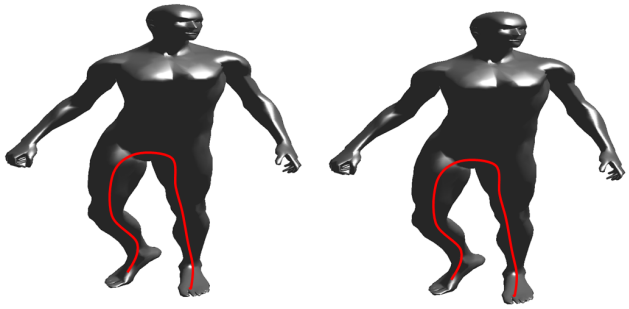

Figure 4.4: Limitation of the proposed symmetry detection algorithm.

which can be further manipulated for other applications as well.

4.3.10 Limitations

The proposed framework fails to detect intrinsic symmetry in cases where the second symmetry criterion (in equation (4.5)) is not satisfied. For example, Figure 4.4, depicts two human figures that constitute what may be perceived as a translational symmetry, although they do not follow our basic input assumption of *connectedness*. The second symmetry criterion (equation (4.5)) fails to hold in this situation. Generally speaking, the proposed algorithm is not designed to detect all forms symmetry resulting from repeated patterns, especially if the patterns are spatially disconnected. In contrast to the approach of Xu et al. [XU12], the use of biharmonic distance instead of geodesic distance as the intrinsic distance measure ensures that the proposed algorithm is capable of detecting intrinsic symmetry even in the presence of small perturbations (such as small bumpy regions) on the 3D surface [LIPMAN10b].

4.4 Symmetry Characterization using Functional Maps

In general, the characterization of a specific transformation based on its functional map is a challenging task. However in our case, the proposed CSV framework ensures that the point pairs used in the generation of the functional map are intrinsically symmetric to a reasonable extent. This property of intrinsic symmetry ensures that the resulting functional map is diagonal or close to diagonal [OVSJANIKOV12]. However, in reality, there are several cases where the actual symmetry transformation deviates substantially from ϵ-isometric deformation resulting in a densely populated functional matrix C. Note that an ϵ-isometric deformation arises from discretization error (ϵ-error) in a perfectly isometric deformation in the continuous case. In the context of the symmetry characterization problem, we assume that the degree of off-diagonality of C corresponds to the complexity of the symmetry transformation, i.e., more non-zero off-diagonal elements in C implies a more complex symmetry transformation. The rationale behind the diagonality assumption is that the CSV-based initialization procedure described in Section 4.3.8 ensures that the two input matrices (representing the Laplace Beltrami eigenfunctions) to the linear equation solver used in the Transformation Space Mapping (TSM) are reasonably similar resulting a diagonal (or close to diagonal) Functional Map matrix C.

The diagonality property of the functional matrix C was first exploited successfully in an intrinsic correspondence framework that used a sparse modeling technique [POKRASS13]. Pokrass et al. [POKRASS13] introduced a weight matrix W of the same dimensions as C where the weight values are assigned using an *inverted* Gaussian model. In the inverted Gaussian model, zero weights are assigned to the diagonal elements whereas the off-diagonal weight values are an increasing function of the distance of the off-diagonal elements from the principal diagonal of the matrix W:

$$w_{i,j} = 1 - \exp(-d_{i,j}^2) \tag{4.6}$$

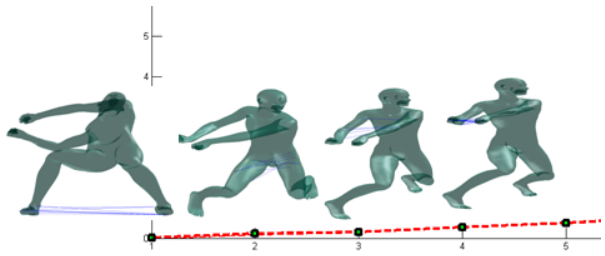

Figure 4.5: Symmetry characterization based on functional maps, in particular, the complexity of the symmetry transformation is characterized by the weight matrix W and represented in the increasing order of the value of the inner product of W and C.

where $d_{i,j}$ is the distance of the $(i,j)^{th}$ element of W from the matrix diagonal.

The complexity of the symmetry transformation is captured by the inner product of C and W given by $\sum_i \sum_j w_{i,j} c_{i,j} = \sum_i \sum_j m_{i,j}$. Since the off-diagonal elements of C are emphasized and the diagonal elements of C are de-emphasized, the inner product can be regarded as a measure of the diagonality of the matrix C. The inner product allows for the successful formulation of a 1D metric symmetry space, wherein each symmetry transformation is represented as a point in the symmetry space with a value given by the inner product of the matrix C and weight matrix W. The Euclidean distance between the points in the 1D symmetry space represents the complexity distance between the corresponding symmetry transformations as depicted in Figure 4.5. In the symmetry space, any perfectly isometric transformation would be mapped onto the origin, whereas more complex symmetry transformations would map onto points farther away from the origin. The symmetry space is a metric space because it possesses all the four primary properties of a distance metric. For two points X and Y in the symmetry space, the distance $d(X,Y)$ from X to Y always follows:

1. *Non-negative definiteness*: $d(X,Y) \geq 0$;

2. *Symmetry*: $d(X, Y) = d(Y, X)$;

3. *Identity of indiscernibles* and *indiscernability of identicals*: $d(X, Y) = 0$ if and only if $X = Y$. Also note that all perfectly isometric deformations map onto the origin of the 1D symmetry space; and

4. *Triangle inequality*: $d(X, Z) \leq d(X, Y) + d(Y, Z)$. Note that although the symmetry transform is a matrix C, we have essentially computed a 1D measure of its diagonality (via the inner product of C and W) to simplify subsequent symmetry analysis. The triangle inequality can be seen to trivially hold in the 1D symmetry space.

It is also possible to cluster the points in the 1D symmetry space to identify intrinsic symmetries which are potentially similar in nature as shown in the Experimental Results section to follow.

4.5 Experimental Results

In this section, we present and discuss the results obtained by the proposed intrinsic symmetry detection algorithm on various 3D shapes. We also provide comparisons of our results with those obtained from the most closely related approaches described in [LIPMAN10a] and [XU12]. We also shown some applications where the detected symmetries can be further analyzed for symmetry characterization and clustering, potentially revealing greater semantic information about the underlying 3D shape. Most of the 3D shape models used in our experiments are from the *Non-rigid World* dataset [BRONSTEIN07] unless mentioned otherwise.

4.5.1 Preprocessing

Several discrete schemes have been proposed in recent years to approximate the Laplace-Beltrami operator on triangular meshes. Among these, the one most widely used for computing the discrete Laplace operator is the *cotangent scheme* (COT), originally proposed by Pinkall and Polthier [PINKALL93]. Belkin et al [BELKIN08] proposed a discrete scheme based on the heat equation, which has been proven to possess the point-wise convergence property for arbitrary meshes. Although it is well known that no discrete Laplace operator can share all of the properties of its continuous counterpart, in our experiments the aforementioned discrete schemes were observed to produce eigenfunctions that approximately preserve the convergence property of the continuous Laplace operator for a reasonably well sampled triangular mesh. In all of the experiments described below, we have used the cotangent scheme [PINKALL93] for the computation of the discrete Laplace-Beltrami operator.

4.5.2 Symmetry Detection

The results of the proposed symmetry detection algorithm are depicted in Figures 4.1, 4.6, 4.7 and 4.11. Several important properties of the proposed algorithm are highlighted in these results.

General Symmetry Detection

The ability of the proposed algorithm to identify multiple intrinsic symmetries is evident from the results shown in Figures 4.6 and 4.11. The extracted symmetries are seen to cover the global symmetry of the underlying 3D shape which has undergone approximate isometric deformations. Additionally, the proposed algorithm is also observed to be capable of detecting symmetry transformations that cover individual components of a 3D object that possess various forms of self-symmetry.

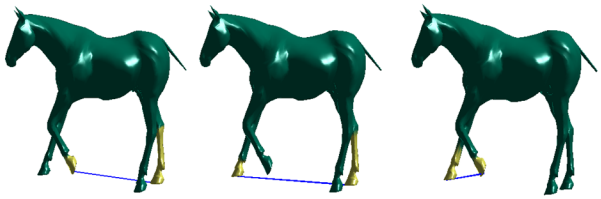

Figure 4.6: Examples of overlapping symmetry detection for the *Horse* shape model.

Overlapping Symmetry Detection

One particularly important aspect of the proposed algorithm is its ability to detect instances of overlapping symmetry. An instance of overlapping symmetry is deemed to occur when a specific region on the surface of the 3D shape is simultaneously subjected to more than one symmetry transformation and, as a result, is symmetric to more than one region on the 3D shape surface. For example, Figure 4.6 shows some examples of overlapping symmetry between legs of the *Horse* shape model successfully detected by the proposed technique. Three different combinations of symmetry transformations between the four legs of the *Horse* model (out of a total of six possibilities) are depicted in Figure 4.6.

Noise

Although the CSV procedure is stochastic in nature, the proposed biharmonic distance-based voting scheme ensures its robustness to noise. In this experiment, each vertex on the input model surface is randomly displaced along the local surface normal where the displacement is modeled by an independent and identically distributed (iid) random variable drawn from a zero-mean Gaussian distribution with a given value of σ (standard deviation). In particular, Figure 4.7 demonstrates the robustness of the proposed symmetry detection technique to different levels of synthetic Gaussian noise (characterized by the standard deviation σ) added

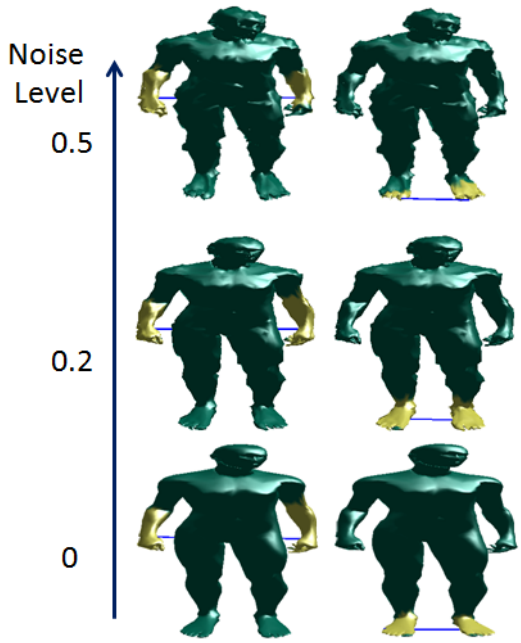

Figure 4.7: Robustness of symmetry detection for three different levels of synthetic white Gaussian noise added to the *Human* shape model.

to the *Human* shape model. Note that the symmetries are consistently detected even at higher noise levels as depicted in Figure 4.7. The details on the generation of the synthetic noisy models are given in Section 4.5.5.

4.5.3 Performance Statistics

All of the experiments reported in this chapter were performed on an Intel Core™ 3.4 GHz machine with 24 GB RAM. For all the example models, the number of sample points were in the range [20, 100]. Table 4.1 reports the timing results for the various steps in the proposed symmetry detection algorithm. In particular, unlike [XU12], wherein the execution time for the most time consuming step of all-pairs geodesic distance computation is not reported, we also report the timing results for the equivalent step in our formulation i.e., the all-pairs

Table 4.1: Timing results for the various steps in the proposed algorithm. Biharmonic: time taken to compute the all-pairs biharmonic distance; FMap: time taken to compututе the functional maps; Extraction: time taken for symmetry extraction; Total: total execution time. All the times are measured in seconds.

# Points	Biharmonic	FMap	Extraction	Total
3400	32	2	5	39
10000	41	3	7	51
50000	237	4	15	256

biharmonic distance computation. The most time consuming step in the proposed algorithm, i.e., the all-pairs biharmonic distance computation, accounts for around 80% of the execution time of the proposed algorithm. More importantly, bypassing the two-step spectral clustering procedure described in [XU12] reduces significantly the computation time of the proposed algorithm.

4.5.4 Comparisons

We have compared the proposed symmetry detection algorithm primarily with methods that could be deemed sufficiently similar, such as ones described in [LIPMAN10a] and [XU12]. The symmetry detection technique of Xu et al. [XU12] can detect overlapping partial intrinsic symmetries whereas that of Lipman et al. [LIPMAN10a] is designed to deal with partial extrinsic symmetries. However, if the symmetric sub-shapes do not undergo significant pose variations, the global alignment component of [LIPMAN10a] may allow it to detect certain partial intrinsic symmetries as well. Though both methods [LIPMAN10a]; [XU12] are capable of *detecting* instances of partial intrinsic symmetry, neither is able to *characterize* the underlying symmetry.

In contrast, the proposed algorithm, not only detects overlapping intrinsic symmetries, but it also has the ability to characterize and cluster the detected symmetries in symme-

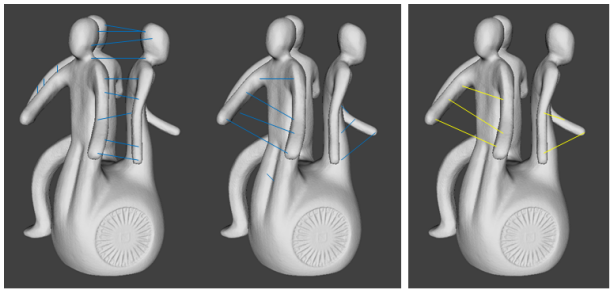

Figure 4.8: The proposed CSV procedure ensures detection of instances of global overlapping intrinsic symmetry (a), (b). In contrast, the symmetry-factored embedding (SFE) technique of Lipman et al. [LIPMAN10a] primarily detects instances of global intrinsic non-overlapping symmetry but fails to detect instances of overlapping symmetry (c).

try space. Since their entire formulation is based solely on global intrinsic distance-based voting followed by a computationally complex two-stage spectral clustering procedure, the technique of Xu et al. [XU12] suffers from the shortcoming of not being scalable in real situations. The proposed algorithm, on the other hand, interpolates the functional map of symmetry transformations from the chosen point pair, to the remaining point pairs, i.e., once a representative point pair within a transformation is identified using the voting procedure, the remaining correspondences are obtained by solving a system of linear equations using the functional map formulation. Although the proposed procedure produces instances of symmetry detection that are qualitatively similar to those in [XU12], its superior scalability is evident even from the timing results reported in Table 4.1 on relatively smaller-size meshes. Moreover, the procedure of Xu et al. [XU12] is unable to produce the underlying maps, thereby circumventing the ultimate goal of characterizing and understanding the intrinsic symmetry.

The symmetry-factored embedding (SFE) technique of Lipman et al. [LIPMAN10a], though designed primarily for extrinsic symmetry detection, in simpler cases, is able to

detect intrinsic symmetries as well. However, it fails completely in detecting cases of overlapping intrinsic symmetries resulting in poor symmetry-based segmentation. In contrast, the proposed CSV procedure ensures the detection of instances of overlapping intrinsic symmetry as shown in Figure 4.8.

4.5.5 Performance in Presence of Noise

It is critically important for a newly proposed symmetry detection technique to perform well in the presence of varying levels of noise. We performed a quantitative evaluation of the proposed technique on the TOSCA dataset [12] under varying levels of additive white Gaussian noise. Each vertex on model surface is randomly displaced along the local surface normal where the displacement is modeled by an independent and identically distributed (iid) random variable drawn from a zero-mean Gaussian distribution with a given value of σ (standard deviation). For each pair of symmetric point correspondences chosen at the end of CSV, the L^2 norm between the values of the original functional map and the noisy functional map is computed and normalized with respect to the original functional map. The mean normalized error in the functional map is computed over all the point correspondences for varying σ values (in the range $[0, 1]$) of the white Gaussian noise as depicted in Figure 4.9. It is interesting to note that the mean normalized error initially increases with increasing values of σ before settling upon a value of ≈ 0.0007 for values of $\sigma > 0.4$, thus depicting the robustness of proposed technique to additive white Gaussian noise.

4.5.6 Quantitative Evaluation

In this section, we present quantitative evaluation of the results obtained by the proposed approach. In order to evaluate the proposed algorithm, we used the SHREC 2010 feature detection and description benchmark [BRONSTEIN10]. This dataset comprises of three

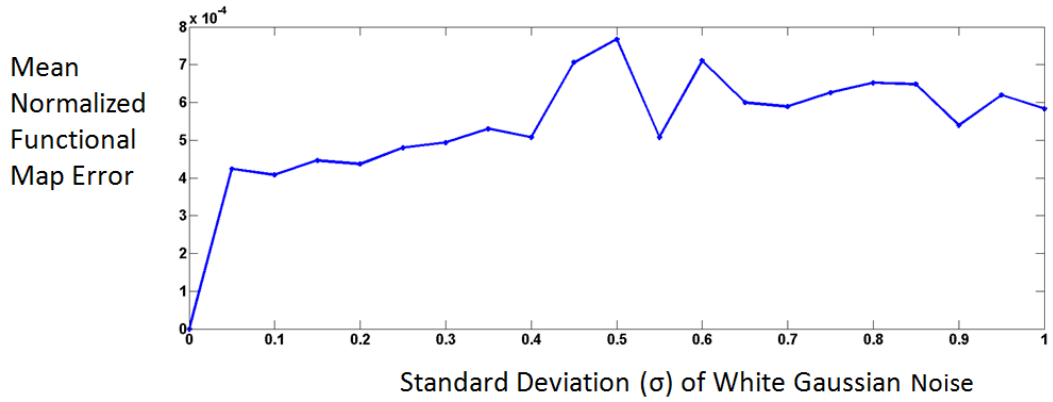

Figure 4.9: Robustness of the proposed technique for varying levels of additive white Gaussian noise.

baseline shapes (i.e., null shapes) and a set of shapes obtained by applying a set of transformations on the null shapes. Shapes have approximately 10,000 to 50,000 vertices and their surfaces are represented by triangular meshes. We have specifically considered shapes which have undergone changes characterized by isometry, topology, micro-holes, scale and noise.

The quantitative evaluation is along the lines proposed by Sipiran and Bustos [SIPIRAN12] and its goal is to determine whether the extracted symmetric components are consistent between a null shape and the transformed shape. Given a null shape X and its transformed shape Y, the extracted symmetric components are represented as $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively. Using the ground-truth data, we compute for each component $Y_j \in Y$ the corresponding component in X which is denoted as X'_j. The component repeatability between X and Y is defined as :

$$R(X, Y) = \sum_{j=1}^{m} \max_{1 \leq i \leq n} O(X_i, X'_j) \tag{4.7}$$

where, the overlap $O(X_i, X'_j)$ between two extracted symmetric components X_i and X'_j is

defined as:

$$O(X_i, X'_j) = \frac{A(X_i \cap X'_j)}{A(X_i \cup X'_j)} \tag{4.8}$$

where $A(X)$ is the surface area of component X.

In addition, it is possible to define the repeatability in overlap O as the percentage of symmetric components in the entire collection $Y_1, \ldots, Y_m$ that have an overlap greater than O with their corresponding components in the null shape [SIPIRAN12]. Note that totally coincident components yield a repeatability value of 1. A plot of repeatability versus overlap can be used to depict the resilience of the symmetry extraction algorithm to various transformations of the null shapes. A resilient symmetry extraction algorithm would be expected to exhibit high values of repeatability (close to 1) for a wide range of overlap values. In our quantitative evaluation, a total 50 sample points were generated using the farthest point random sampling strategy for each of the null shapes and transformed shapes. Figure 4.10 shows the plot of repeatability versus overlap for various transformations such as isometry, changes in topology, introduction of micro-holes, addition of noise and changes in scale. As can be observed in Figure 4.10, the proposed symmetry extraction algorithm maintains a high overall repeatability O ($\geq 80\%$) at overlap values ≤ 0.75 for all transformations except for changes in topology. The reason for lower repeatability over the range of overlap values in case of topological changes, is that topological changes have the potential to alter significantly the intrinsic distance (i.e., the geodesic distance and the biharmonic distance) measurements.

4.5.7 Symmetry Characterization

The diagonality property of the functional map could be used to characterize the underlying symmetry transformation and classify it as a simple transformation or as one that is more complex in nature. Our assumption is that, greater the complexity of symmetry transfor-

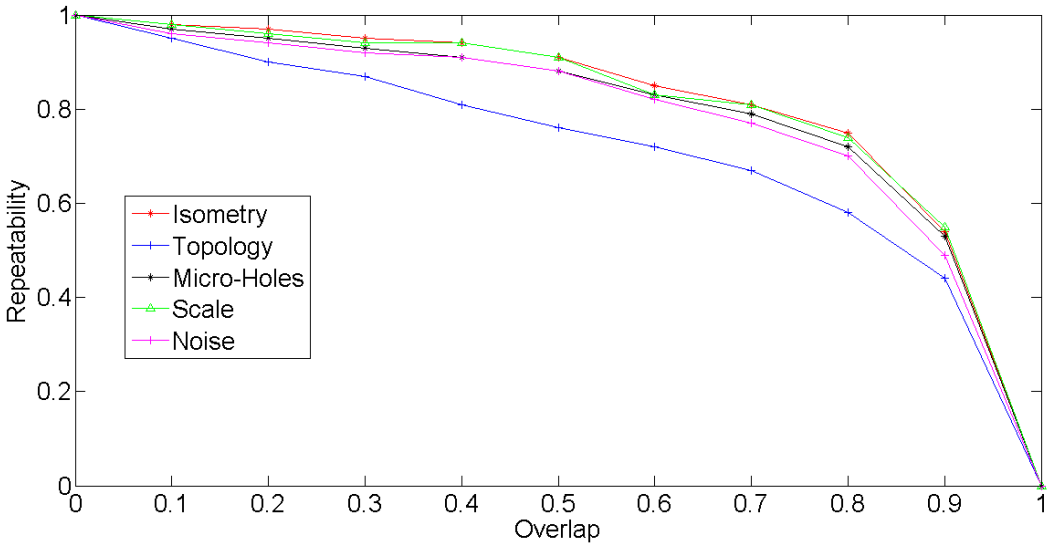

Figure 4.10: Plot of overlap vs. repeatability for the proposed symmetry detection technique on the SHREC 2010 benchmark dataset.

mation, greater the deviation of the shape deformation from intrinsic isometry, resulting in a deformation characterized by a functional matrix C with higher off-diagonal element values which results in a higher value for the inner product of the functional matrix C and weight matrix W. The resulting characterization of the isometric deformation is depicted in Figure 4.5.

4.5.8 Symmetry Group Retrieval

As discussed in [XU12], symmetry has a group structure. Consequently, the retrieval of symmetry transformation can be cast as a clustering problem from an algorithmic perspective. We exploit the functional maps generated previously to cluster the detected instances of symmetry in transformation space and retrieve the symmetry groups. The CSV procedure is known to inherently possess the symmetric flip problem where the directions denoting the correspondence between the points in a point pair could be reversed. This problem is re-

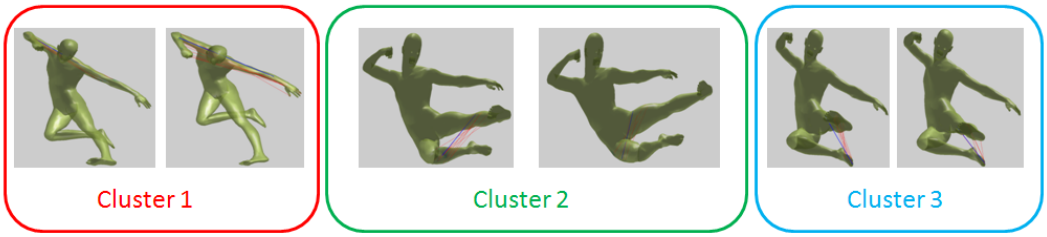

Figure 4.11: Clustering of the functional maps with representative correspondences resulting in three symmetry groups.

solved via adjusting the relative positions of source and destination points within each point pair by finding the minimum distance between the source and destination points in the bi-harmonic distance space. This ensures that the functional maps generated from potentially similar transformations will have similar structure since the symmetry flip problem between the maps is resolved. In particular, we have used a simple k-means clustering algorithm to cluster the functional maps based on their symmetry groups as depicted in Figure 4.11.

4.6 Conclusion and Future Directions

We have presented an algorithm for detection and characterization of intrinsic symmetry in 3D shapes. While the results obtained are encouraging, we regard our work as an initial attempt towards the larger goal of achieving complete understanding of and a verifiable solution to the general problem of symmetry detection and characterization. We have identified some limitations of our approach and we hope to address these in our future work.

To the best of our knowledge, this is one of the first attempts to formalize the symmetry analysis problem not only as one of symmetry detection, but as one that can be extended to include symmetry characterization and symmetry clustering in the transformation space. In

particular, the introduction of the functional map formalism in symmetry detection enables us to come up with a novel representation of the symmetry transformation as a map. In future, we aim to formulate operations, such as addition and subtraction, on these generated maps that would potentially provide a deeper and more comprehensive understanding of intrinsic symmetry in general.

The incorporation of *Transformation Space Mapping* in symmetry characterization is a novel idea and its full potential can only be realized after more extensive experimentation. In particular, we plan to study the possibility of map based exploration of similar symmetry transformations across shapes in the near future. Another important direction that can be considered is the possibility of incorporating this technique in *Computer-Aided Geometric Design* (CAGD) for urban architecture. In urban architecture, symmetric repetition of the same pattern is a common phenomenon. During the design phase, if the basic pattern is stored explicitly only once and the symmetric repetitions are saved as functional maps, the resulting reduction in space- and time-complexity of the design problem would be significant.

Bibliography

[AUBRY11] AUBRY, M., ULRICH, S., and DANIEL, C. (2011) The wave kernel signature: A quantum mechanical approach to shape analysis. *Proc. Computer Vision Workshops (ICCV Workshops)*.

[BOKELOH09] BOKELOH, M., BERNER, A., WAND, M., SEIDEL, H.-P., and SCHILLING, A. (2009) Symmetry detection using line features. *Computer Graphics Forum (Eurographics)* 28(2), pp. 697–706.

[BRONSTEIN10] BRONSTEIN, A. M., BRONSTEIN, M. M., BUSTOS, B., CASTELLANI, U., CRISANI, M., FALCIDIENO, B., GUIBAS, L. J., KOKKINOS, I., MURINO, V., SIPIRAN, I., OVSJANIKOV, M., PATANE, G., SPAGNUOLO, M., and SUN, J. (2010) SHREC 2010: Robust feature detection and description benchmark. *Proc. Workshop on 3D Object Retrieval (3DOR10)*.

[BRONSTEIN07] BRONSTEIN, A. M., BRONSTEIN, M. M., and KIMMEL, R. (2007) Calculus of non-rigid surfaces for geometry and texture manipulation. *IEEE Trans. Visualization and Computer Graphics*, Vol 13(5), September-October, pp. 902–913.

[BRONSTEIN08] BRONSTEIN, A. M., BRONSTEIN, M. M., and KIMMEL, R., "Numerical geometry of non-rigid shapes", Springer, 2008. ISBN: 978-0-387-73300-5.

[BELKIN08] BELKIN, M., SUN, J., and YUSU, W. (2008) Discrete Laplace operator on meshed surface. *Proc. Symp. Comp. Geometry*, pp. 278–287.

[BERNER08] BERNER, A., BOKELOH, M., WAND, M., SCHILLING, A., and SEIDEL, H.-P. (2008) A graph-based approach to symmetry detection. *Proc. Fifth Eurographics/IEEE VGTC Conf. Point-Based Graphics*, Eurographics Association, pp. 1–8.

[ELDAR97] ELDAR, Y., LINDENBAUM, M., PORAT, M., and ZEEVI, Y. Y. (1997). The farthest point strategy for progressive image sampling. *IEEE Transactions on Image Processing*, 6(9), pp. 1305–1315.

[IMIYA99] IMIYA, A., UENO, T., and FERMIN, I. (1999). Symmetry detection by random sampling and voting process. *Proc. IEEE Conf. Image Analysis and Processing (CAIP99)*, pp. 400–405.

[KAZDHAN03] KAZHDAN, M., CHAZELLE, B., DOBKIN, D., FUNKHOUSER, T., and RUSINKIEWICZ, S. (2003) A reflective symmetry descriptor for 3D models. *Algorithmica*, 38(1), October, pp. 201–225.

[LI05] LI, W., ZHANG, A., and KLEEMAN, L. (2005) Fast global reflectional symmetry detection for robotic grasping and visual tracking. *Proc. Australasian Conference on Robotics and Automation (ACRA05)*, December.

[LI06] LI, M., LANGBEIN, F., and MARTIN, R. (2006) Constructing regularity feature trees for solid models. *Geometric Modeling and Processing-GMP 2006*, Springer Berlin Heidelberg, pp. 267–286.

[LING06] LING, H., and OKADA, K. (2006) Diffusion distance for histogram comparison, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2006)* , Vol. 1, pp. 246–253.

[LIU07] LIU, S., MARTIN, R., LANGBEIN, F., and ROSIN, P. (2007) Segmenting periodic reliefs on triangle meshes. *Mathematics of Surfaces XII*, Springer Berlin Heidelberg, pp. 290–306.

[LIPMAN10a] LIPMAN, Y., CHEN, X., DAUBECHIES, I., and FUNKHOUSER, T. (2010). Symmetry factored embedding and distance. *ACM Transactions on Graphics (TOG)*, 29(4), pp. 103.

[LIPMAN10b] LIPMAN, Y., RUSTAMOV, R. M., AND FUNKHOUSER, T. A. (2010). Biharmonic distance. *ACM Transactions on Graphics (TOG)*, 29(3), pp. 27.

[LITMAN13] LITMAN, R., and BRONSTEIN, A. (2013) Learning spectral descriptors for deformable shape correspondence. *IEEE Trans. Pattern Analysis and Machine Intelligence*, in press.

[LORENSEN87] LORENSEN, W.E., and CLINE, H.E. (1987) Marching Cubes: A high resolution 3D surface reconstruction algorithm. *ACM SIGGRAPH Computer Graphics*, 21(4), pp. 163–169.

[MITRA06] MITRA, N. J., GUIBAS, L. J., and PAULY, M. (2006) Partial and approximate symmetry detection for 3D geometry. *ACM Transactions on Graphics (TOG)*, 25(3), pp. 560–568.

[MITRA13] MITRA, N. J., PAULY, M., WAND, M., and CEYLAN, D. (2013) Symmetry in 3D geometry: Extraction and applications. *Computer Graphics Forum*, Blackwell Publishing Ltd.

[OVSJANIKOV08] OVSJANIKOV, M., SUN, J., and GUIBAS, L. (2008) Global intrinsic symmetries of shapes. *Computer Graphics Forum* 27(5), Blackwell Publishing Ltd., pp. 1341–1348.

[OVSJANIKOV10] OVSJANIKOV, M., QUENTIN, M., MEMOLI, F., and GUIBAS, L. J. (2010) "One point isometric matching with the heat kernel." *Computer Graphics Forum*, 29(5), pp. 1555-1564.

[OVSJANIKOV12] OVSJANIKOV, M., BEN-CHEN, M., SOLOMON, J., BUTSCHER, A., and GUIBAS, L. (2012). Functional maps: A flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)*, 31(4), pp. 30.

[PAULY08] PAULY, M., MITRA, N. J., WALLNER, J., POTTMANN, H., and GUIBAS, L. (2008) Discovering structural regularity in 3D geometry. *ACM Transactions on Graphics (TOG)*, 27(3), pp. 43:1–11.

[PINKALL93] PINKALL, U., and POLTHIER, K. (1993) Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics*, 2(1), pp. 15–36.

[PODOLAK06] PODOLAK, J., SHILANE, P., GOLOVINSKIY, A., RUSINKIEWICZ, S., and FUNKHOUSER, T. (2006) A planar-reflective symmetry transform for 3D shapes. *ACM Transactions on Graphics (TOG)* 25(3), July, pp. 549–559.

[POKRASS13] POKRASS, J., BRONSTEIN, A. M., BRONSTEIN, M. M., SPRECHMANN, P., and SAPIRO, G. (2013) Sparse modeling of intrinsic correspondences. *Computer Graphics Forum*, 32(2.4), Blackwell Publishing Ltd., pp. 459–468.

[RAVIV10] RAVIV, D., BRONSTEIN, A. M., BRONSTEIN, M. M., and KIMMEL, R. (2010) Full and partial symmetries of non-rigid shapes. *Intl. Jour. Computer Vision* 89(1), pp. 18-39.

[RAVIV07] RAVIV, D., BRONSTEIN, A., BRONSTEIN, M., and KIMMEL, R. (2007) Symmetries of non-rigid shapes. *Proc. Intl. Conf. Computer Vision*, pp. 1–7.

[RUSTAMOV08] RUSTAMOV, R. (2008) Augmented planar reflective symmetry transform. *The Visual Computer* 24(6), pp. 423-433.

[SCHRODINGER26] SCHRODINGER, E. (1926) An undulatory theory of the mechanics of atoms and molecules. *Physical Review* 28(6), pp. 1049-1070.

[SHIKHARE01] SHIKHARE, D., BHAKAR, S., and MUDUR, S. (2001) Compression of large 3D engineering models using automatic discovery of repeating geometric features. *Signal Processing*, 19(20), pp. 233-240.

[SIPIRAN12] SIPIRAN, I., and BUSTOS, B. (2012). Key-component detection on 3D meshes using local features. In Proceedings of the 5th *Eurographics conference on 3D Object Retrieval*, pp. 25–32.

[WANG11] WANG, Y., XU, K., LI, J., ZHANG, H., SHAMIR, A., LIU, L., CHENG, Z., and XIONG, Y. (2011) Symmetry hierarchy of man-made objects. *Computer Graphics Forum*, 30(2), Blackwell Publishing Ltd., pp. 287-296

[XU09] XU, K., ZHANG, H., TAGLIASACCHI, A., LIU, L., LI, G., MENG, M., and XIONG, Y. (2009) Partial intrinsic reflectional symmetry of 3D shapes. *ACM Transactions on Graphics (TOG)*, 28(5), pp. 138:1-10.

[XU12] XU, K., ZHANG, H., JIANG, W., DYER, R., CHENG, Z., LIU, L., and CHEN, B. (2012) Multi-scale partial intrinsic symmetry detection. *ACM Transactions on Graphics (TOG)*, 31(6), pp. 181.

[YEN07] YEN, L., FOUSS, F., DECAESTECKER, C., FRANCQ, P., and SAERENS, M. (2007) Graph nodes clustering based on the commute-time kernel. *Proc. 11th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2007)*, Springer Lecture notes in Computer Science, LNCS.

[YIP00] YIP, R. (2000) A Hough transform technique for the detection of reflectional symmetry and skew-symmetry. *Pattern Recognition Letters*, 21(2), p. 117–130.

[ZABRODSKY95] ZABRODSKY, H., PELEG, S., and AVNIR, D. (1995) Symmetry as a Continuous Feature, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(12), pp. 1154–116.

Chapter 5

A 3D BoF Descriptor for Analysis of the Left Ventricular endocardial Surface Morphology

5.1 Introduction

Since *Coronary Artery Disease* (CAD) is a leading cause of morbidity and mortality worldwide [7], techniques that increase the effectiveness and/or lower the costs of diagnostic or prognostic procedures associated with CAD are expected to have a significant clinical impact. CAD is caused by *atherosclerosis* or accumulation of lipoprotein plaque in the coronary arteries that supply blood to the *myocardium* or cardiac muscle tissue. Atherosclerotic plaques lead to the progressive narrowing or *stenosis* of the coronary arteries, resulting in reduced blood flow and consequently, reduced oxygen supply to the myocardium, a condition termed as *myocardial ischemia*. If untreated, myocardial ischemia may lead to irreversible necrosis of the myocardium wherein the healthy myocardium is increasingly replaced by scar tissue, thus compromising the cardiac function and resulting in congestive cardiac failure. Some plaques

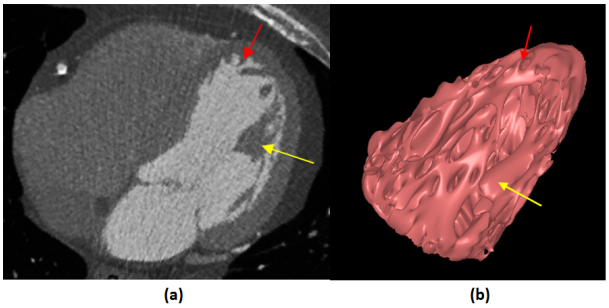

Figure 5.1: Left ventricular trabeculae carneae (indicated by the red arrow) and papillary muscles (indicated by the yellow arrow) can be clearly seen in the raw CTCA images (a) acquired using a 320-MDCT scanner and in the reconstructed triangular mesh (b).

that are deemed vulnerable may suddenly rupture resulting in coronary artery occlusion and cardiac arrest, potentially leading to sudden death.

X-Ray Coronary Angiography (XRA) is an invasive technique that is a clinically accepted standard for assessment of vascular morphology and the extent of vessel stenosis due to artherosclerotic plaque deposition. However, a *comprehensive* assessment of CAD necessitates a study of both, vascular morphology and cardiovascular function. Conventional cardiovascular functional assessment is performed via a stress-induced perfusion test that uses Magnetic Resonance (MR) or radionuclide Myocardial Perfusion Imaging (MPI). Since vascular morphology and cardiovascular function are imaged using separate modalities, the time and cost associated with a comprehensive assessment of CAD and the potential health risk to the patient associated with higher radiation doses are all significantly increased.

Computed Tomography Coronary Angiography (CTCA) is a non-invasive imaging technique that allows for robust quantification of vascular morphology and has the potential for characterizing the atherosclerotic plaque composition [20]. When performed using a 320-Multirow-Detector Computed Tomography (MDCT) scanner, a CTCA scan can be performed within a single heart beat, yielding images with an isotropic spatial resolution of 0.5

mm in a volumetric fashion. The resulting images, in addition to providing vascular morphology information, are capable of providing significant details about the left ventricular (LV) endocardial surface structure. The CTCA images, backed by anatomical studies, have revealed that, instead of a simple and smooth surface, the LV endocardial surface is composed of a complex structure of *trabeculae carneae*, which are small muscular columns that arise naturally from the inner surface of the ventricles (Figure 6.1). Also revealed in the CTCA images are the *papillary* muscles which are attached to the cusps of the atrioventricular (i.e., the mitral) valves (Figure 6.1).

Alterations in the LV trabeculation structure have been clinically observed to closely associate with some cardiovascular diseases, such as myocardial noncompaction disease [1] and hypertrophy [7]. However, due to the limitations in the spatial resolution of conventional *in vivo* imaging techniques, very few research studies have been undertaken to formally investigate the LV trabeculation structure at a detailed level, and quantify the relationships between structural changes in LV trabeculation, changes in cardiovascular function and clinically observed cardiovascular pathologies. If the analysis of the LV endocardial surface structure can be shown to provide significant insights into vascular morphology, cardiovascular function and the progression of CAD then it has the potential to provide significant additional or supplementary diagnostic value to the results of MPI, CTCA and XRA. In some cases, it could potentially reduce the need for invasive and/or stress-based testing procedures that are expensive, time consuming and pose a greater health risk to the patient, thus making CTCA imaging using MDCT scanners, a potential one-stop-shop technique for assessment of both, vascular morphology and cardiovascular function.

The sub-endocardial layer of the myocardium is the first and the most susceptible cardiac region in the development of CAD on account of its inherently higher oxygen consumption requirement and the restricted collateral blood flow resulting in reduced oxygen delivery during the progression of CAD. The presence of intermediate to severe CAD in patients is often

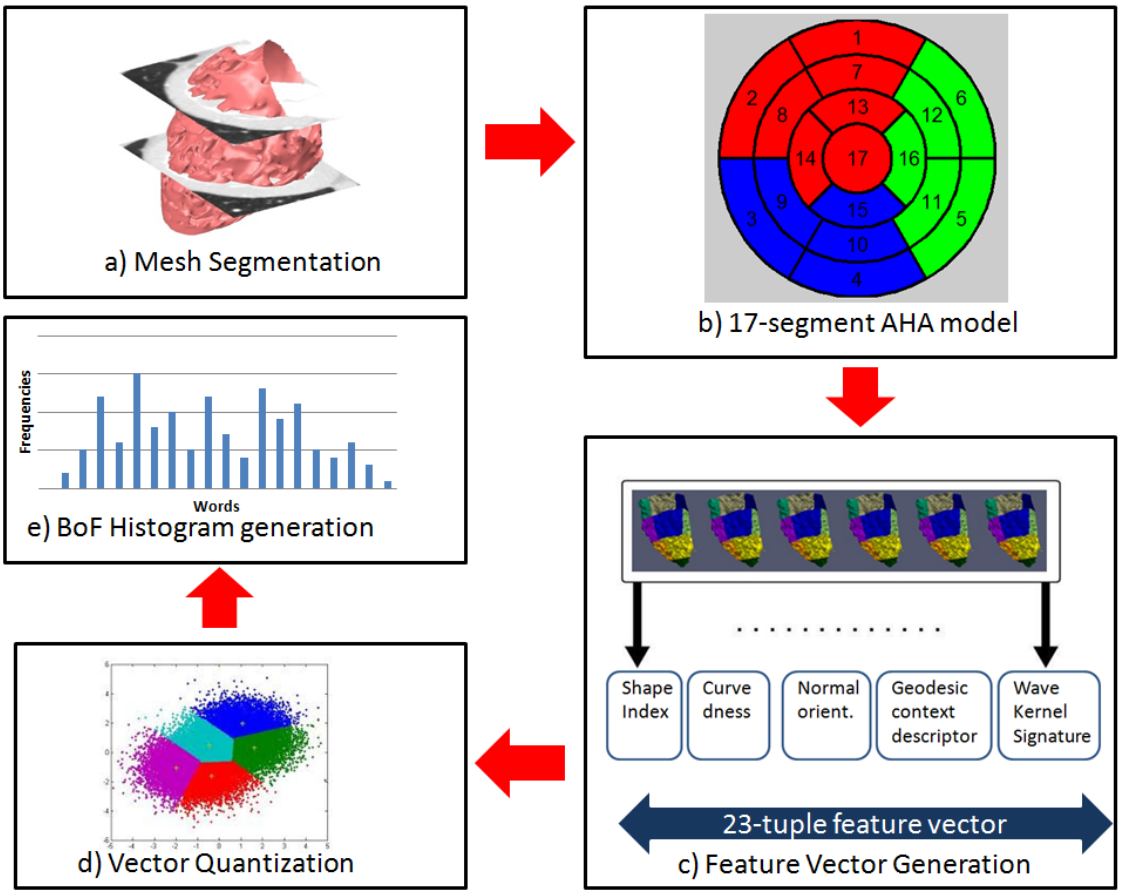

Figure 5.2: Illustration of the sequence of steps in the morphological analysis of the LV endocardial surface: (a) accurate mesh segmentation followed by (b) generation of a 17-segment LV surface model with demarcation of coronary arterial territories (red: LAD, green: LCX, blue: RCA), (c) feature vector generation and, (d, e) final generation of the BoF histogram via vector quantization (k-means clustering).

clinically suggestive of varying degrees of myocardial ischemia in the corresponding regions of the myocardium. Furthermore, the affected ischemic myocardium loses contractibility and has a tendency to get stiffer and be pushed outward by the high LV blood pressure. These myocardial changes are reflected in the LV endocardial surface morphology, in that the structure of the trabeculae and the papillary muscles on the LV endocardial surface are clinically observed to be substantially altered [19]. This clinically hypothesized association between the LV endocardial surface morphology and the incidence and severity of CAD constitutes the central thesis of this chapter.

The objective of this chapter is to formally establish the clinically observed relationship between the incidence and severity of CAD, as quantified by the extent and location(s) of coronary arterial stenosis, and the structural alterations in the LV endocardial surface as observed in high-resolution MDCT images. The LV endocardial surface structure, specifically the structure of the LV endocardial trabeculae and papillary muscles, is formally characterized using 3D morphological features obtained via 3D shape analysis algorithms. To the best of our knowledge, this research is amongst the first to formally characterize the LV endocardial surface structure and quantify the clinically observed association between the alterations in the LV endocardial surface structure and the extent and location(s) of coronary arterial stenosis. The research described in this chapter could potentially add significant value to the accuracy of diagnosis and effective management of CAD. It constitutes a first step towards formal investigation, quantification and verification of the clinically observed association between the LV endocardial surface morphology, cardiovascular function and vascular morphology which, in the long term, would aid in comprehensive assessment and understanding of the physiological mechanisms underlying the development and progression of CAD.

5.2 Related Work

Our preliminary work [19] demonstrated the potential diagnostic value of characterization of the LV endocardial surface structure in assessing the incidence of CAD. Although it produced some significant and encouraging results, our preliminary work did exhibit shortcomings that stemmed primarily from the selection of shape descriptors used to characterize the LV endocardial surface. The two shape descriptors that were employed were based on the implicit assumption of rigidity of the LV endocardial surface as observed in the MDCT images. Consequently, these shape descriptors were observed to have significant limitations in terms of their classification accuracy [19]. Although the MDCT image data were collected at a relatively steady phase of 75% in the R-R cardiac cycle, the continuous LV motion and the human error involved in the data collection, tested severely the limits of the rigidity assumption. It was clear that a more robust shape descriptor was needed to characterize the LV endocardial surface, i.e., one that is invariant to isometric global shape deformation [15, 28].

Isometry-invariant shape descriptors have been the topic of recent research, especially in content-based retrieval in shape databases [22], shape symmetry detection [18], dense surface correspondence determination [28] and surface registration [23]. However, their use in the characterization of the LV endocardial surface has not been explored in the research literature. In this chapter we show that the incorporation of isometry-invariant features in the LV endocardial surface descriptor enables more accurate and robust characterization of the underlying LV endocardial surface.

Deriving a feature-based representation of the contents of an image generally entails two stages of processing, *feature detection* and *feature description* [22]. The goal of feature detection is to localize relatively stable points or regions within an image that possess significant information and can be repeatedly and reliably detected in transformed versions of the im-

age. Different approaches are employed for feature detection based on the intrinsic nature of the features, i.e., point-based or region-based, and the desired scale of abstraction for the features.

Region-based feature detection techniques typically rely on segmentation of the cardiac images using shape priors [6, 24]. The high-dimensional shape prior is projected into a low-dimensional subspace using dimensionality reduction techniques, e.g., variants of Principal Component Analysis (PCA), to constrain the shape variation. The detected features are tracked in a cardiac image sequence using a Kalman filter [6] or particle filter [24] whereas the shape prior is adapted using a learning algorithm such as AdaBoost [24]. Other region-based feature detection techniques for cardiac images include model-based deformable templates [5], Markov random fields [16], optical flow techniques [12] and combinations of the above [2].

The goal of the feature description stage is to arrive at a representation of the local image information in the neighborhood of the detected feature. In order to achieve a parsimonious representation, a feature vocabulary consisting of *visual words* is constructed by performing vector quantization in the feature descriptor space using a clustering technique such as the k-means clustering algorithm. After having performed vector quantization, the individual feature descriptors are replaced by indices in the vocabulary of visual words. The aggregation of feature descriptors to describe the overall shape is achieved by generating a frequency histogram of the visual words in the vocabulary, termed as a *Bag-of-Features* (BoF).

One of the prominent implementations of the BoF concept is *Video Google*, a web-based application designed by Sivic and Zisserman [26, 27] for object-based search in large image and video collections. Likewise, two shapes are compared using their respective visual word frequency histograms or BoF representations. This reduces the shape similarity problem to the problem of frequency histogram comparison. *Shape Google*, a technique for shape-based search in large image collections developed by Ovsjanikov et. al. [22] and, the shape comparison approach proposed by Toldo et. al. [29] are amongst the prominent works on

BoF-based shape description and shape comparison, and the work described in this chapter is influenced by both of these works [22, 29].

As an extension to the general BoF approach, we propose a novel BoF-based shape analysis approach designed specifically for cardiac image analysis, and for medical image analysis in general. The proposed BoF-based shape analysis approach is designed to quantify the relationship between the incidence, severity and localization of CAD and the structural alterations in the LV endocardial surface. While several works have proposed feature-based approaches for characterization of rigid 3D shapes, very few are capable of dealing with non-rigid 3D shape deformations [17]. To the best of our knowledge, this is one of the first works that uses a BoF-based shape analysis approach for comparison of non-rigid deformable 3D shapes in the context of cardiovascular imaging in particular, and medical imaging in general.

In this chapter, a BoF-based approach is proposed to encapsulate the local and global geometry, local surface orientation and global contextual information for the LV endocardial surface. The proposed approach is shown to result in a robust feature vector for the purpose of morphological analysis of the LV endocardial surface. As shown in the experimental results, the proposed approach results in successful localization of coronary arterial stenosis which serves to strengthen the clinically observed relationship between the incidence and severity of CAD and morphological alterations in the LV endocardial surface. The sequence of steps in the proposed approach for morphological analysis of the LV endocardial surface is depicted in Figure 5.2.

The remainder of the chapter is organized as follows: In Section 5.3, the main contributions of the chapter are outlined; in Section 6.3, the proposed LV surface segmentation and LV shape analysis procedures are detailed; and in Section 5.5, experimental results are presented. Finally, in Section 5.6, the chapter is concluded with a brief discussion of the proposed approach and an outline of directions for future work.

5.3 Contributions of the chapter

The chapter makes two principal contributions as described below:

1. The *Bag-of-Features* (BoF) framework for non-rigid shape analysis is adapted for the purpose of cardiac shape analysis which is an important problem in cardiac imaging in particular and medical imaging in general. It is also important to note that the proposed approach is sufficiently general to be applicable to other problems in medical imaging such as 3D shape analysis of the human brain and 3D shape-based search and retrieval in large medical databases.

2. The chapter proposes a geometric and machine learning-based model of the relationship between localized changes in the LV endocardial surface morphology and the incidence and extent of stenosis in specific coronary arteries. This constitutes an important initial step towards clinical understanding of the complex relationship between coronary arterial stenosis and its effect on the morphology of the LV endocardial surface. To the best of our knowledge, this chapter represents one of the first attempts to model this complex clinical relationship in a mathematically structured manner.

5.4 MDCT Image Segmentation and LV Shape Analysis

5.4.1 LV endocardial surface segmentation and meshing

The *trabeculae carneae* along the LV endocardial surface can be broadly classified into three different morphological types: (a) those that lie along the entire length of the LV wall forming prominent ridges; (b) those that are fixed at their extremities but free in the middle; and (c) those that connect the roots of the papillary muscles to the ventricular wall. These

different *trabeculae* morphologies contribute to the complexity of the LV endocardial surface topology [14].

A 3D level set approach is employed to segment the LV endocardial surface while adapting to the topological changes caused by the complex trabeculation structure [10]. A median filter-based denoising procedure is employed on the 3D MDCT data prior to segmentation in order to suppress noise while retaining the edges in the MDCT images. Unless mentioned otherwise, the size of the median filter is empirically set to 7×7 based on the MDCT data set. A level set-based segmentation procedure without reinitialization, as proposed by Li et al. [10], is applied to the median-filtered 3D image data set followed by the *Marching Cubes* procedure [11] to generate the surface meshes. The surface meshes are subsequently denoised via a mean face normal filtering procedure proposed by Zhang and Hamza [31] to obtain the smooth shape of the LV myocardial surface.

5.4.2 Data preparation

The standardized myocardial segmentation model proposed by the American Heart Association (AHA) is adopted to enable accurate segmentation, understanding and localization of cardiac anatomy and pathology [14]. The AHA-approved 17-segment cardiac model [4] is adapted to divide the left ventricle into 17 segments for more accurately localized shape analysis (Figure 5.3). The long axis of the left ventricle is first computed to divide the LV endocardial surface into 4 main segments along the longitudinal orientation termed as the *apex, apical, mid-cavity* and *basal*. Further division of the LV endocardial surface along the short axis view is tackled by exploiting knowledge of the cardiac anatomy. Three landmark points are considered across the septum to perform the division process. This results in the division of the *apical* segment into 4 segments and the *mid-cavity* and *basal* segments each into 6 segments whereas the *apex* segment remains undivided. At the end of the segmentation process, the LV endocardial surface is divided into 17 segments in conformity to the

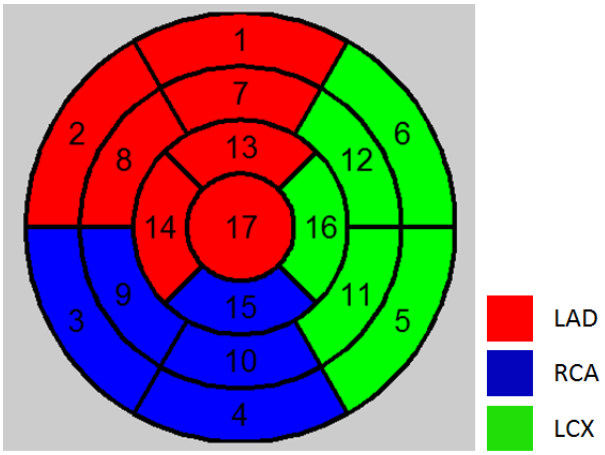

Figure 5.3: Bull's eye view of the 17-segment AHA model. Each LV segment is color coded to denote the territory of the major coronary artery it falls under, i.e., Left Anterior Descending (LAD): red; Right Coronary Artery (RCA): blue; and Left Circumflex Artery (LCX): green.

AHA model as shown in Figure 5.3.

The ground truth data for our experimental validation is obtained via an XRA procedure performed on the same subjects. XRA is the clinical standard for the diagnosis of coronary artery stenosis. Note that XRA is an invasive imaging modality distinct from non-invasive MDCT. In particular, the XRA procedure results in the percent diameter stenosis (%DS) data for each major coronary artery.

5.4.3 LV endocardial surface descriptors based on the rigidity assumption

In our preliminary work [19], we considered two primary shape descriptors to characterize the LV endocardial surface morphology, i.e., the $D2$ *shape descriptor*, a global shape descriptor proposed by Osada et al. [21] and the *shape index*, a local shape descriptor first introduced by Koenderink [9] and subsequently modified by Zaharia and Preteux [30]. Both descriptors

make an implicit assumption of rigidity of the endocardial surface and are invariant to scale and 3D rigid body transformation (i.e., rotation and translation) in Euclidean space.

The $D2$ shape descriptor is a global shape signature of a 3D object and is represented by a probability distribution function. The probability distribution function is obtained via sampling of a pre-specified local shape function which measures a basic geometric property of the underlying 3D shape. In the case of the $D2$ shape descriptor, the shape function is the distance between two randomly sampled vertices on the mesh-based representation of the 3D LV surface. Osada et al. [21] have shown that the simple $D2$ shape descriptor, when modeled as a probability distribution function over the entire 3D object surface, can serve as an effective object signature for a variety of applications such as 3D shape-based retrieval from a 3D object database.

The *shape index*, in contrast, describes the local shape of the 3D surface mesh based on the values of the principal surface curvatures computed within a local neighborhood of a given 3D surface point. The shape index is particularly effective in describing the *local extrinsic* geometry of the surface. The shape index I_p of a surface point p is a function of the two principal surface curvatures $\kappa_1(p)$ and $\kappa_2(p)$ associated with point p and is defined as follows:

$$I_p = \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\kappa_1(p) + \kappa_2(p)}{\kappa_1(p) - \kappa_2(p)}\right) \tag{5.1}$$

where, $\kappa_1(p) > \kappa_2(p)$.

The value of the shape index lies within the interval [0, 1] and enables surface-based representation of basic elementary shapes such as *convex, concave, rut, ridge* and *saddle* [30]. Since the global structure and the detailed local structure of the LV endocardial surface are both potentially clinically relevant and since there have been no previous studies detailing which of the shape properties (local or global) convey more valuable clinical information, we

considered both local and global shape descriptors in our preliminary work [19]. However, the results of our preliminary work showed the shape index to be more effective than the $D2$ shape descriptor for classification of the sample LV endocardial surface data into *normal* and *diseased* classes.

5.4.4 LV endocardial surface descriptors based on non-rigid deformation

BoF-based approaches for shape representation typically consist of two steps, i.e., feature detection and feature description. For representation of object shapes encountered in common imagery, the feature detectors and feature descriptors are typically designed to be invariant under affine transformation to take into account the different possible viewpoints under which a 3D object may appear in an image. For non-rigid characterization of the LV endocardial surface, the type of invariance required is different from the one required when dealing with object shapes in common images since the number of surface transformations in the case of the latter is much larger, including bending, shearing, changes in pose and changes in connectivity. In our case, the LV endocardial surface deformation can be constrained to lie within a pre-specified range since all the CTCA images are acquired at a relatively steady phase (75%) within the R-R cardiac cycle. The aforementioned steady phase assumption allows the LV surface deformation to be approximated by an *isometric* deformation. Consequently, the LV endocardial surface descriptors need to be robust enough to capture the important characteristics of the underlying shape while exhibiting invariance to isometric deformation. Since the LV endocardial surface is typically less rich in terms of features than common images, the task of detecting a large number of stable, repeatable and isometric deformation-invariant feature points is significantly harder. In this work, the features selected for description of the LV endocardial surface are inspired by the work of

Toldo et. al. [29].

Feature detection

Although the feature detection step is not mandatory for BoF-based shape description, it can improve computational efficiency in practice. In our case, the feature detection strategy comprises of dense random sampling. We consider as feature points, 500 randomly sampled surface points from each of the 17 LV endocardial surface segments. The underlying motivation is to ensure that the sampling is as dense as possible so that the entire spectrum of interesting points on the LV endocardial surface is considered for further analysis. In situations where availability of computational power is not a constraint, the feature selection step can be ignored and all the points on the surface can be considered for the subsequent feature description step.

Feature description

The feature descriptors are designed to ensure that both the local and global geometric properties of the surface are represented in sufficient detail. Inspired by the work of Toldo et. al. [29], four surface descriptors are designed to represent each sampled surface point. The first three descriptors represent the local geometry of the surface around a specific point whereas the fourth is a contextual descriptor which encapsulates the global properties of a specific point with respect to other points on the surface. The descriptors are detailed below:

Shape Index (I_p) The *shape index* I_p of a surface point p, is computed using equation (5.1) as previously described. The shape index is particularly important for its ability to quantitatively characterize the local shape of a surface.

Curvedness (C_p) The *curvedness* descriptor represents the degree of local curvature of a surface. Normalized *curvedness* also encapsulates the rate of change of local surface curva-

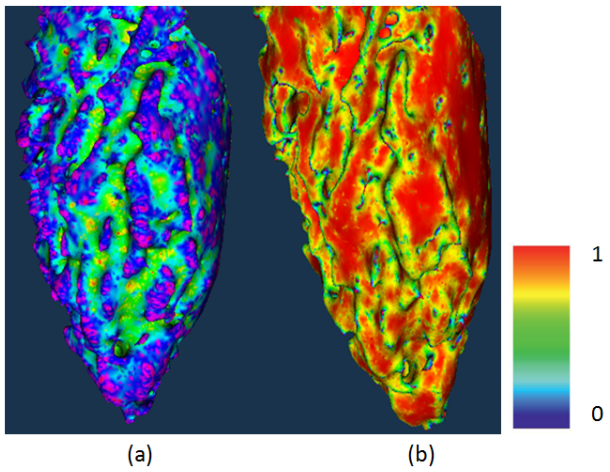

Figure 5.4: Illustration of the relationship between the (a) shape index I_p and (b) curvedness C_p surface descriptors for a sample LV endocardial surface.

ture: a value close to 1 implies a very gentle change, whereas a value close to 0 implies a very sharp change. Generally, points on a smooth spherical object have small *curvedness* values. The *curvedness* C_p at a surface point p, as proposed by Koenderink [9], is defined as a function of the two principal surface curvatures $\kappa_1(p)$ and $\kappa_2(p)$ associated with surface point p as follows:

$$C_p = \sqrt{\frac{(\kappa_1(p))^2 + (\kappa_2(p))^2}{2}} \tag{5.2}$$

where, $\kappa_1(p) > \kappa_2(p)$. The relationship between the two surface descriptors, i.e., the shape index I_p and the curvedness C_p for a typical LV endocardial surface is depicted in Figure 5.4. Like the shape index I_p, the value of curvedness C_p is also invariant to scale and 3D rigid-body transformation (i.e., translation and rotation) in Euclidean space.

Normal Orientation (θ_p) It has been clinically observed that the incidence of myocardial infarction and myocardial ischemia, can cause changes in the 3D orientation of the LV

trabeculation structure due to the build up of scar tissue. Consequently, the local 3D orientation of the LV endocardial surface is incorporated within the feature vector. The *normal orientation* θ_p at surface point p is defined as the angle between the unit normal vector at surface point p and the XZ-plane.

Geodesic Contextual Descriptor ($GCD_p^{f_1}$) The *Geodesic Contextual Descriptor* (GCD) is intended to provide a means for describing the overall shape and measuring shape similarity based on point correspondences. Contextual information is used to describe a specific surface point p with respect to the 3D shape as a whole. The geodesic contextual descriptor for a surface point p encapsulates the distribution of relative positions of other points on the same surface with respect to p, thus summarizing global shape in an informative and, most importantly, in an isometric deformation-invariant manner.

Since BoF-based methods typically employ a frequency distribution histogram, there is inherent loss of contextual information. However, semantically distinct results can be disambiguated only via incorporation of contextual information. Since it is difficult to incorporate contextual information within a frequency distribution histogram, it needs to be incorporated within the feature vector. The *geodesic contextual descriptor*, denoted by $GCD_p^{f_1}$, is an isometric deformation-invariant contextual descriptor that encapsulates the relative position of a surface point p in relation to the other points on the surface segment. It is characterized by a f_1-bin histogram which is generated by computing the normalized geodesic distance between the surface point p and other sampled points on the surface segment. The geodesic contextual descriptor $GCD_p^{f_1}$ is invariant to scale, 3D rigid body transformation and isometric deformation.

Wave Kernel Signature-based Descriptor ($WKS_p^{f_2}$) The comprehensive analysis of shapes entails the formulation of a geometric feature descriptor which characterizes each point on the object surface while taking into account its relation to the entire shape. To

determine the *Wave Kernel Signature* (WKS) of a surface point x, one evaluates the probability of a quantum particle with a certain energy distribution to be located at point x. The behavior of the quantum particle on the surface is governed by the Schrodinger equation [25]. Assuming that the quantum particle has an initial energy distributed around some nominal energy with a probability density function $f(e)$, the solution of the Schrodinger equation can then be expressed in the spectral domain as:

$$\psi(x,t) = \sum_{k \geq 1} \exp(ie_k t) f(e_k) \phi_k(x) \tag{5.3}$$

Aubry et al. [3] considered a family of log-normal energy distributions centered around a mean log energy $\log e$ with variance σ^2. This particular choice of distributions is motivated by a perturbation analysis of the Laplacian spectrum [3]. Having fixed the family of energy distributions, each point x on the surface is associated with a WKS of the form: $p(x) = (p_{e_1}(x), \ldots, p_{e_n}(x))^T$ where $p_{e_i}(x)$ is the probability of measuring a quantum particle with the initial energy distribution $e_i(x)$ at point x. Aubry et al. [3] use logarithmic sampling to generate the values $e_1(x), \ldots, e_n(x)$.

Since the WKS is derived from the Laplace-Beltrami operator, it is inherently invariant under isometric deformation. Additionally, the WKS exhibits a band-pass characteristic. This reduces the influence of low frequencies and allows for better separation of frequency bands across the descriptor dimensions. Since the WKS is a multi-scale signature we considered the WKS at f_2 different values of scale for each point. The WKS-based descriptor at point p is formally denoted as $WKS_p^{f_2}$.

The feature detection and feature description procedures described above result in a N-tuple feature vector for each surface point p denoted by $F_p = (I_p, C_p, \theta_p, GCD_p^{f_1}, WKS_p^{f_2})$ as depicted in Figure 5.2, where $N = f_1 + f_2 + 3$.

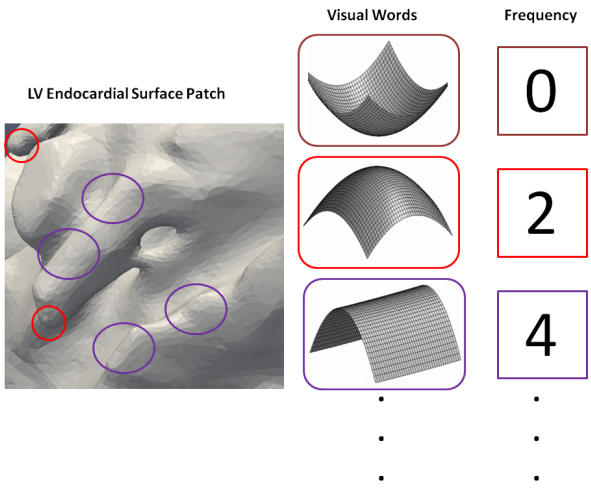

Figure 5.5: Concept visualization for the BoF histogram generation procedure. The frequency distribution histogram of the sampled *visual words* is illustrated for a sample LV endocardial surface.

Construction of visual vocabularies

The feature vectors F_p at each surface point are clustered in order to obtain the *visual words*. Assuming that the local descriptors are computed for a set of stable surface points, the feature vector space is quantized to obtain a compact representation for the vocabulary of visual words, in a manner similar to the Shape Google approach [22]. A *vocabulary* of visual words is defined as a set of representative vectors in the descriptor space or feature space, obtained by means of unsupervised learning, i.e., vector quantization via k-means clustering in our case. More formally, a *vocabulary* of visual words is defined as a collection $V = \{v_1, \ldots, v_k\}$ where v_i is the centroid of the ith cluster and the clusters represent the visual words. We employed an adaptive k-means clustering algorithm for generating the final histogram. Figure 5.5 depicts the frequency distribution histogram of the sampled visual words for a sample LV endocardial surface.

5.5 Experimental Results

5.5.1 Experimental setup

The proposed methods for segmentation, meshing and endocardial surface shape description were tested on 32 MDCT data sets consisting of 16 data sets from cardiac patients and 16 data sets from normal subjects. The incidence of single- or multi-vessel obstructive CAD was found in the three major coronary arteries, i.e., the Left Anterior Descending Artery (LAD), Left Circumflex Artery (LCX) and Right Coronary Artery (RCA) using XRA, and further confirmed by Myocardial Perfusion Imaging (MPI) and Fractional Flow Reserve (FFR) tests performed on the patients. A detailed description of the ground truth data generated via the XRA procedure for all 32 subjects is shown in Table 5.1. The individual territories for each of the three major coronary arteries are depicted in Figure 5.3 in conformity to the AHA convention.

Table 5.1: Ground Truth Data generated for all the 32 subjects via the XRA procedure. LAD: Left Anterior Descending Artery, LCX: Left Circumflex Artery, RCA: Right Coronary Artery. The columns corresponding to LAD, LCX and RCA denote the locations and extent of stenoses in the corresponding coronary arteries; Prox: proximal, Mid: mid portion, Dis: distal, Dx: diagonal branches, PDA: Posterior Descending Coronary Artery, Ostium: Origin of the coronary artery just above the aortic valve, Normal: unconstricted coronary artery.

Case	Disease	LAD	LCX	RCA
01 - 16	Normal			
17	Diseased	Mid: 60% D1:50% D2 prox:80% D2 dis:99%	Normal	Prox 70-75%
18	Diseased	Mid: 30-40%	40% towards the origin of the PDA	Prox:40% Mid:60-70%
19	Diseased	Mid: 70%	Prox: 40-50%	Mid: 30-40%
20	Diseased	Mid: 95%	Prox: 80% Ostium: 90%	Prox: 20%
21	Diseased	Prox: 40%	Prox: 75%	Prox: 30% Dis: 40% PDA: 20-30%
22	Diseased	Mid: 70%	Normal	Normal
23	Diseased	Normal	Prox: 90%	Borderline
24	Diseased	Normal	80%	Normal
25	Diseased	90%	30%	Normal
26	Diseased	60-70%	40-50%	80%
27	Diseased	80%	Normal	Normal
28	Diseased	Normal	Normal	90-95%
29	Diseased	90%	90%	40%
30	Diseased	70%	50-60%	40%
31	Diseased	95%	Normal	Normal
32	Diseased	20%	80%	40-50%

The cardiac patients and normal subjects were subject to a contrast-enhanced CTCA scan on a 320-MDCT scanner using a standard protocol with electrocardiogram (ECG) gating. The resulting images were reconstructed at a relatively steady state of 75% in the R-R cardiac cycle to ensure minimal ventricular motion. Although the above protocol is designed to ensure that the endocardial surfaces are reconstructed at a relatively fixed cardiac dilation stage and that the subsequent shape analysis is minimally affected by cardiac motion, it still results in the presence of motion artifacts in the acquired images due to various factors such as the high velocity of the ventricular motion and irregular ECG. However, the assumption of isometric deformation enables us to better account for these motion artifacts. The segmentation technique described in Section 5.4.1 was used to generate topologically correct and geometrically accurate data for subsequent analysis.

The experimental results are detailed for three different scenarios. In the first scenario, the problem is modeled as one of *global* classification wherein the shape index (I_p) histogram, the rigid-body shape descriptor (I_p, C_p, θ_p) histogram, the combined rigid-body and isometric deformation-invariant shape descriptor ($I_p, C_p, \theta_p, GCD_p^{f_1}, WKS_p^{f_2}$) histogram and the BoF histogram are each used to classify the entire LV endocardial surface into one of two classes, i.e., *normal* and *diseased*. In the second scenario, the BoF-based approach is used to model the problem as one of *localized* classification where each LV segment is labeled as *normal* or *diseased* based on the extent of diameter stenosis (DS) in the major coronary arteries supplying blood to the LV segment under consideration. The LV segment is considered *normal* if the percent DS (%DS) value of the corresponding major coronary artery is less than 70% and *diseased* if it is $\geq$ 70%. In the third scenario, the BoF-based approach is used to model the problem as one of multivariate regression where the exact %DS values of each of the major coronary arteries are treated as the labels based on which the correlation coefficients for each of the LV segments are computed.

5.5.2 Segmentation results

The results of the LV endocardial surface segmentation have already been proven to be reasonably accurate in our preliminary study [19]. The spatial distribution of the trabeculation was observed by clinicians to vary with the location within the left ventricle; thus providing the rationale for using the standard 17-segment AHA model to perform localized shape analysis. Furthermore, the proposed segmentation approach has already shown a visually observable distinction in trabeculation between normal and diseased hearts, yielding a classification accuracy greater than 80% with simple rigid-body surface descriptors and a nearest-neighbor classifier [19]. Our previous work has demonstrated the accuracy of the proposed segmentation method as well as its applicability for subsequent quantitative shape analysis [19]. In this particular experiment, we varied the size of median filter kernel and considered three different sizes, i.e., 5×5, 7×7 and 9×9 to determine the effect of the median filter size on the final surface mesh.

Table 5.2: Quantitative evaluation of the segmentation framework: Average volume overlap of the segmented region with respect to the ground truth for different size of the median filter kernel.

Median Filter Kernel Size	5×5	7×7	9×9
Average Volume Overlap	0.796	0.925	0.831

In order to optimize the filter kernel and validate the segmentation method, an *in vitro* experiment was conducted to study the correlation between the segmentation accuracy and the image noise level. An excised porcine heart was put in a plastic container and immersed in the CT contrast solution. The air in the ventricle was carefully squeezed out and completely removed. The porcine heart was scanned with different levels of tube current to emulate the different radiation exposure conditions that result in CT images of varying quality.

The detailed CT imaging parameters include: volumetric acquisition; tube voltage: 120 kVp; tube current: 50-550 mA; and reconstruction slice thickness: 0.5 mm. The CT images

were acquired with 550 mA tube current, which is the highest possible tube current setting available on the CT scanner, and used as the ground truth. Compared to the human chest in clinical studies, the very small volume of the excised porcine heart and the highest possible tube current setting ensured adequate image quality for the purpose of ground truth validation.

The *in vitro* images of the porcine heart acquired for varying tube current values in the range 50 mA - 400 mA were used for validation of the segmentation procedure, since the image noise levels were observed to be in a range similar to that of clinical *in vivo* CT images. Quantitative evaluation of the segmentation procedure was performed by tuning the median filter kernel size for the CT images of the porcine heart. For each tube current setting, median filtering was performed for three different kernel sizes, i.e., 5×5, 7×7 and 9×9. The volume overlap V_O between the segmented volume S_V and the volume generated by the ground truth segmentation S_G was computed as:

$$V_O = \frac{(S_V \cap S_G)}{(S_V \cup S_G)} \tag{5.4}$$

The average V_O results over all tube current setting values for varying values of the median filter kernel size are reported in Table 5.2. Based on the results in Table 5.2, for all subsequent experiments, the median filter kernel size was fixed at 7×7 since it was observed to provide a good balance between smoothness and preservation of geometric details of the LV endocardial surface.

5.5.3 Parameter Estimation

The main three parameters need to be chosen for our method are f_1 (number of bins in the GCD histogram), f_2 (number of discrete scales of the WKS-based descriptor) and k (number of clusters in the k-means clustering algorithm). We computed the mean classification ac-

curacy using an artificial neural network (ANN), employing a multilayer perceptron (MLP) architecture with a single hidden layer and a learning rate of 0.3, over all the 17 LV segments across all 32 subjects for different combinations of values for the parameters f_1, f_2 and k and reported the results in Table 6.1. The mean classification accuracy denotes the fraction of LV segments across all 32 subjects that are correctly classified as *normal* or *diseased*. Since the combination of parameter values $f_1 = 10$, $f_2 = 10$ and $k = 20$ was observed to yield the best classification accuracy, we used this set of parameter values for all the experiments reported in the remainder of the chapter. Thus, the N-tuple feature vector described earlier is essentially a 23-tuple feature vector in all of our experiments.

Table 5.3: Parameter tuning for maximizing the mean classification accuracy over all the 17-segments across 32 subjects.

GCD (f_1)	WKS (f_2)	k	Mean Classification Accuracy
10	10	10	0.51
10	10	20	0.57
10	20	10	0.51
10	20	20	0.49
20	10	10	0.49
20	10	20	0.53
20	20	10	0.48
20	20	20	0.54

5.5.4 Global classification results

As the first part of global classification experiment, we tested a series of classifiers using weka [8] with several combinations of respective parameters. These classifiers include decision trees, random forest, linear SVMs, SVMs with radial basis function kernel and Artificial Neural Networks.

The experiment suggested that the artificial neural network (ANN), employing a multilayer perceptron (MLP) architecture with a single hidden layer and a learning rate of 0.3 was

doing the best classification, and was used as a result, for the purpose of global classification in a manner similar to that described in [20]. The ANN-based classification procedure was carried out within a strictly leave-one-out setting. The Shape Index, Curvedness, Normal Orientation (I_p, C_p, θ_p)-based histogram was observed not to result in any improvement in classification over the Shape Index I_p-based histogram (Tables 5.4 and 5.5).

Table 5.4: Confusion matrix to illustrate the classification accuracy of the Shape Index I_p-based histogram.

	Classified Diseased	Classified Normal
Actual Diseased	13	3
Actual Normal	3	13

Table 5.5: Confusion matrix to illustrate the classification accuracy of the Shape Index, Curvedness, Normal Orientation (I_p, C_p, θ_p)-based histogram.

	Classified Diseased	Classified Normal
Actual Diseased	13	3
Actual Normal	3	13

In the case of the $(I_p, C_p, \theta_p, GCD_p^{10}, WKS_p^{10})$-based histogram, 27 out of 32 samples were classified correctly resulting in an overall accuracy of 84.37%. Table 5.6 shows the confusion matrix for the diagnosis accuracy. The false alarm rate was observed to be 18.75% whereas the miss rate was observed to be 12.5%. Thus, the incorporation of the 10-bin Geodesic Context Descriptor GCD_p^{10} and 10-scale Wave Kernel Signature WKS_p^{10} in the feature vector was observed to improve the overall classification accuracy on account of its invariance to isometric deformation.

To demonstrate the superiority of the BoF-based description, the binary classification experiment described above was repeated for the BoF frequency histogram. In this experiment, 29 out of 32 samples were classified correctly resulting in an overall accuracy of 90.62%. Table 5.7 shows the confusion matrix for the diagnosis accuracy. The false alarm rate was

Table 5.6: Confusion matrix to illustrate the classification accuracy of the 23-tuple feature vector $(I_p, C_p, \theta_p, GCD_p^{10}, WKS_p^{10})$-based histogram.

	Classified Diseased	Classified Normal
Actual Diseased	14	2
Actual Normal	3	13

observed to be 6.25% whereas the miss rate was observed to be 12.5%.

Table 5.7: Confusion matrix to illustrate the classification accuracy of the BoF histogram.

	Classified Diseased	Classified Normal
Actual Diseased	14	2
Actual Normal	1	15

5.5.5 Localized classification results

A series of experiments was performed to demonstrate the effectiveness and limitations of the non-rigid shape descriptor $F_p = (I_p, C_p, \theta_p, GCD_p^{10}, WKS_p^{10})$. In this series of experiments, a coronary artery was considered as *diseased* or *stenotic* if the extent of stenosis was 70% or greater. The LV myocardial segments were labeled as diseased if they belonged to the stenotic artery's territory in conformity to the AHA convention. The available %DS data were used for determining whether a coronary artery is *normal* or *diseased*. The 20-bin BoF frequency histograms (Section 5.5.3) resulting from the vector quantization procedure (Section 5.4.4), for a particular LV segment from all the LV datasets across the 32 subjects, were used as the inputs to the MLP ANN with same configuration parameters as the one described in Section 5.5.4. Figure 5.6 illustrates the discriminative power of the BoF frequency histograms in being able to distinguish between normal and diseased subjects. The output of the MLP ANN is the classification of a particular LV segment as *normal* or *diseased*.

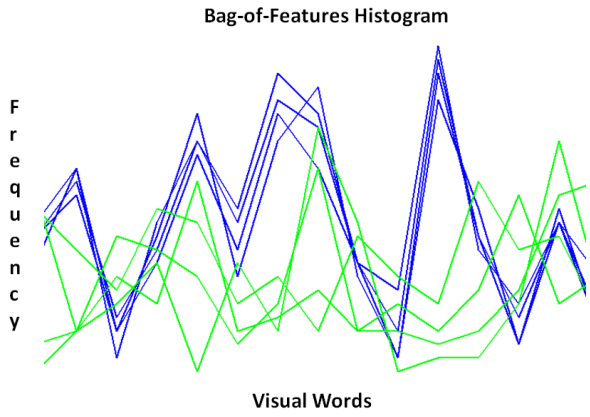

Figure 5.6: Illustration of a sample BoF frequency histogram which demonstrates its discriminative power. The normal subjects (blue) exhibit a frequency histogram pattern that is very distinct from that of the diseased subjects (green). A particular region of the BoF frequency histogram is shown here to depict the difference between diseased subjects and normal subjects in this particular region. This, in turn, is reflected in the higher classification accuracy for the BoF histogram.

The success rate for detection of stenosis for a specific coronary artery is tabulated in Table 5.8 and also depicted in Figure 5.7. The classification results depict a clinically observed relationship between the coronary arterial stenosis and the affected segment in the 17-segment AHA model. The lower classification accuracy in the basal area (segments 1-6) can be explained by the clinical observation that several instances of coronary arterial stenosis encountered in this study are located in the mid to distal portion of the coronary arteries that only affect the mid-cavity (segments 7-12) and apical (segments 13-16) portions of the LV endocardial surface. Furthermore, another probable reason for the lower classification rate in the basal area is that the apical and mid-cavity segments exhibit greater endocardial trabeculation structure than the basal segments, which translates to more reliable endocardial surface morphology information that can be used for the purpose of classification in the case of the apical and mid-cavity segments.

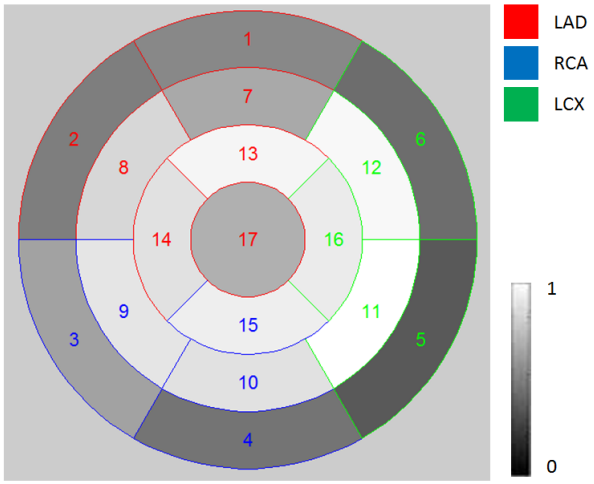

Figure 5.7: Illustration of the classification accuracy for detection of coronary arterial stenosis based on the change of surface morphology in the 17 LV segments in the AHA model. Higher gray values denotes higher classification accuracy and vice versa. The numerical label for each LV segment is color coded to denote the territory of the major coronary artery it falls under: LAD: red; RCA: blue; and LCX: green.

5.5.6 Non-rigid shape descriptor-based localized multivariate regression results

In this experiment, the classification problem of Section 5.5.5 was recast as a multivariate linear regression problem wherein the exact %DS data was considered as the ground truth. The regression procedure was carried out within a strictly leave-one-out setting. Multivariate linear regression attempts to fit a linear model to multiple independent (or explanatory) variables $x_1, x_2, \ldots x_n$, to obtain an estimate for a single dependent (or response) variable y as shown below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots + \beta_n x_n + \epsilon \tag{5.5}$$

where the β_i's are the model coefficients and ϵ is the fitting error. In our case, the dependent variable y denotes the exact %DS value and the independent variables x_i, $i = 1, \ldots, n = 20$

Table 5.8: Classification accuracy, Pearson's correlation coefficient (PCC) and coefficient of determination R^2 for multivariate regression on a per AHA-defined LV segment basis.

Coronary Artery	Segment Number	Classification Accuracy	PCC	R^2
LAD	1	0.41	0.32	0.1
	2	0.38	0.27	0.07
	7	0.51	0.76	0.57
	8	0.65	0.55	0.3
	13	0.74	0.62	0.38
	14	0.68	0.64	0.4
	17	0.53	0.42	0.17
RCA	3	0.49	0.29	0.08
	4	0.35	0.25	0.06
	9	0.69	0.61	0.37
	10	0.68	0.63	0.39
	15	0.72	0.72	0.51
LCX	5	0.27	0.31	0.09
	6	0.33	0.29	0.08
	11	0.77	0.74	0.54
	12	0.75	0.77	0.59
	16	0.71	0.59	0.35

denote the components of the 20-tuple BoF histogram feature vector generated via the vector quantization procedure described in Section 5.4.4.

During the training process, the exact %DS data for a major coronary artery (LAD, LCX and RCA) and the 20-tuple BoF histogram feature vectors corresponding to the AHA-defined LV segments that comprise the territory of that coronary artery are fed as inputs to the multivariate linear regression procedure. The multivariate linear regression procedure determines the model coefficients, i.e., β_i's, that minimize the mean squared error given by $\frac{1}{m}\Sigma_{k=1}^{m}\epsilon_m^2$ where m is the number of exact %DS value points associated with the major coronary artery (equation (5.5)).

During the testing procedure, the *estimated* %DS value, y_{est}, for the major coronary is

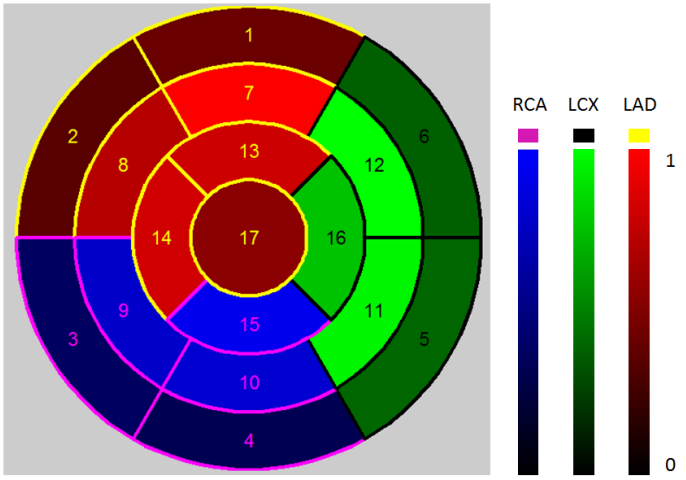

Figure 5.8: Illustration of the correlation coefficient for detection of coronary arterial stenosis based on the change in surface morphology in the 17 LV segments in the AHA model. For each coronary arterial territory a separate color is used to denote the correlation coefficient value, e.g., LAD: red; RCA: blue; and LCX: green. The numerical label for each LV segment is color coded to denote the territory of the major coronary artery it falls under: LAD: yellow; RCA: magenta; and LCX: black.

computed using the 20-tuple BoF histogram feature vector $(x_1, x_2, \ldots, x_{20})$ and the coefficients $(\beta_0, \beta_1, \ldots, \beta_{20})$ as follows:

$$y_{est} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots + \beta_n x_n \tag{5.6}$$

The Pearson's correlation coefficient (PCC) and the coefficient of determination (R^2) between the estimated outcome y_{est} of equation (5.6) and the exact %DS data y (i.e., the ground truth) are computed.

The values of the PCC and R^2 for changes in the LV morphology (as characterized by the 20-tuple BoF histogram feature vectors) with respect to the %DS values in a localized LV segment are tabulated in Table 5.8 and depicted in Figure 5.8. The regression results are seen to depict an intriguing and clinically observed relationship between the extent of

coronary arterial stenosis and the affected LV segment in the 17-segment AHA model. The lower PCC and R^2 values in the basal area (segments 1-6) compared to those in the mid-cavity (segments 7-12) and apical portions (segments 13-16) of the LV endocardial surface could be explained in a manner similar to that described in Section 5.5.5.

5.5.7 Non-rigid shape descriptor-based coronary arterial territory-wise classification and multivariate regression results

The relatively low values of the classification accuracy and correlation coefficients resulting from the proposed BoF-based approach might stem from the use of the LV territories defined on the basis of the current AHA convention. However, the AHA convention may not accurately reflect the true anatomy in individual patient cases. Consequently, we repeated the experiments described in Sections 5.5.5 and 5.5.6, where instead of focusing on the individual AHA-defined segments that comprise the LV territory of each major coronary artery, we considered the entire LV territory of each major coronary artery in our analysis. Thus, instead of the 17-segment AHA LV model we considered a 3-supersegment LV model where the supersegments comprise the territories of the major coronary arteries LAD, LCX and RCA. The procedures for feature detection, feature selection, classification and regression are the exactly same, except that the shape being considered is not that of one specific segment within the 17-segment AHA LV model but the aggregate shape of all segments comprising the territory of a specific coronary artery, i.e., LAD, LCX or RCA. Since the size of the coronary arterial territory is significantly larger that that of a typical AHA-defined segment, the number of randomly sampled points per coronary arterial territory is 2000 instead of 500 per AHA-defined segment. The results in Table 5.9 show significantly higher values for the classification accuracy and regression correlation coefficient when considered on the basis of each coronary arterial territory compared to the results in Table 5.8 which are on a per

AHA-defined LV segment basis.

Table 5.9: Classification accuracy, Pearson's correlation coefficient (PCC) and coefficient of determination R^2 for multivariate regression when considered on the basis of each coronary arterial territory.

	LAD	LCX	RCA
Classification Accuracy	0.88	0.91	0.92
PCC	0.85	0.90	0.89
R^2	0.72	0.81	0.79

5.6 Conclusions and Future Work

To the best of our knowledge, this chapter is one of the earliest works that studies the relationship between coronary artery stenosis and the morphological alterations in the LV endocardial surface using high-resolution MDCT data, and demonstrates its potential predictive value for the diagnosis of the incidence and severity of CAD. One limitation of our work is the use of the LV territories defined on the basis of the current American Heart Association (AHA) convention. However, the AHA convention may not accurately reflect the true anatomy in individual patient cases which in turn can be an explanation for the low classification accuracy in the localized classification results (Section 5.5.5).

Our investigation also sheds new light on the localization of LV regions that are the most affected by coronary arterial stenosis, a phenomenon which is yet to be fully explained. The association between the morphological features of the LV endocardial surface and cardiovascular function will be explored in our future work. In particular, we aim to investigate the correlation between the LV endocardial surface morphology and the results of Myocardial Perfusion Imaging (MPI) and Fractional Flow Reserve (FFR) tests in addition to the coronary arterial stenosis results obtained via X-Ray Coronary Angiography (XRA).

Bibliography

[1] Y. Agmon, H.M. Connoll, L.J. Olson, B.K. Khandheria, and J.B. Seward, Noncompaction of the ventricular myocardium. *Jour. American Soc. Echocardiography*, Vol. 12(10), Oct. 1999, pp. 859-863.

[2] K. Althoff, G. Hamarneh, and T. Gustavsson, Tracking contrast in echocardiography by a combined snake and optical flow technique, *IEEE Computers in Cardiology*, Vol. 3653, Cambridge, MA, Sept. 2000, pp. 29-32.

[3] M. Aubry., S. Ulrich, and C. Daniel, The wave kernel signature: A quantum mechanical approach to shape analysis. *Proc. Computer Vision Workshops (ICCV Workshops)*, 2011.

[4] M. D. Cerqueira, N. J. Weissman, V. Dilsizian, A. K. Jacobs, S. Kaul, W. K. Laskey, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart, *Circulation*, Vol. 105, 2002, pp. 539-542.

[5] V. Chalana, D.T. Linker, D.R. Haynor, and Y. Kim, A multiple active contour model for cardiac boundary detection on echocardiographic sequences, *IEEE Trans. Med. Imaging* (TMI), Vol. 15(3), 1996, pp. 290-298.

[6] D. Comaniciu, X-S. Zhou, and S. Krishnan, Robust real-time myocardial border tracking for echocardiography: an information fusion approach. *IEEE Trans. Med. Imaging* (TMI), Vol. 23(7), 2004, pp. 849-860.

[7] S. Goo, P. Joshi, G. Sand, D. Gerneke, A. Taberner, Q. Dollie, I. LeGrice, and D. Loiselle, Trabeculae carneae as models of the ventricular walls: implications for the delivery of oxygen, *Jour. Gen. Physiology*, Vol. 134(4), Oct. 2009, pp. 339-350.

[8] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, Vol. 11(1), 2009, pp. 10-18.

[9] J. Koenderink, *Solid Shape*, The MIT Press, Cambridge, Massachusetts, 1990.

[10] C. Li, C. Xu, C. Gui, and M.D. Fox, Level set evolution without re-initialization: a new variational formulation, *Proc. IEEE Conf. CVPR*, Vol. 1, 2005, pp. 430-436.

[11] W. E. Lorensen and H. E. Cline, Marching cubes: A high resolution 3D surface construction algorithm, *Computer Graphics*, Vol. 21(4), 1987, pp. 163169.

[12] G.E. Mailloux, F. Langlois, P.Y. Simard and M. Bertrand, Restoration of the velocity field of the heart from two-dimensional echocardiograms, *IEEE Trans. Med. Imaging*, Vol. 8(2), 1989, pp. 143-153.

[13] J. Matas, O. Chum, M. Urban, and T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, *Image and Vision Computing* Vol. 22(10), 2004, pp. 761-767.

[14] Medtronic Inc, *The Visible Heart* webpage: http://www.visibleheart.com/index.shtml

[15] F. Memoli and G. Sapiro, A theoretical and computational framework for isometry invariant recognition of point cloud data, *Found. Comput. Math.* Vol. 5(3), 2005, pp. 313-347.

[16] M. Mignotte, J. Meunier and J.C. Tardif, Endocardial boundary estimation and tracking in echocardiographic images using deformable templates and Markov random fields, *Pattern Analysis and Applications*, Vol. 4(4), Nov. 2001, pp. 256-271.

[17] N.J. Mitra, L.J. Guibas, J. Giesen, and M. Pauly. Probabilistic fingerprints for shapes, *Proc. Symp. Geometry Processing*, 2005, Vienna, Austria.

[18] N.J. Mitra, L.J. Guibas, M. Pauly. Partial and approximate symmetry detection for 3D geometry, *Proc. ACM SIGGRAPH*, 2006, pp. 560-568.

[19] A. Mukhopadhyay, Z. Qian, S.M. Bhandarkar, T. Liu and S. Voros, Shape analysis of the left ventricular endocardial surface and its application in detecting coronary artery disease, *Proc. Intl. Conf. Funct. Imaging and Modeling of Heart (FIMH)*, 2011, New York, NY.

[20] A. Mukhopadhyay, Z. Qian, S.M. Bhandarkar, T. Liu, S. Rienhart, and S. Voros, Morphological analysis of the left ventricular endocardial surface and its clinical implications, *Proc. MICCAI* , Nice France, Oct. 2012.

[21] A R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, Shape distributions, *ACM Trans. on Graphics*, Vol. 21(4), pp. 807-832, October 2002.

[22] M. Ovsjanikov, A.M. Bronstein, M.M. Bronstein and L.J. Guibas, Shape Google: A computer vision approach to invariant shape retrieval, *Proc. NORDIA Workshop*, 2009.

[23] M. Ovsjanikov, Q. Mrigot, F. Mmoli, and L.J. Guibas, One point isometric matching with the heat kernel, *Eurographics Symposium on Geometry Processing*, 2010

[24] Z. Qian, D. Metaxas, and L. Axel, Boosting and nonparametric based tracking of tagged MRI cardiac boundaries, *Proc. MICCAI*, LNCS 4190, 2006, pp. 636-644.

[25] E. Schrodinger, An undulatory theory of the mechanics of atoms and molecules, *Physical Review*, Vol. 28 (6), 1926, pp. 10491070.

[26] J. Sivic, and A. Zisserman, Video Google: A text retrieval approach to object matching, *Proc. IEEE Intl. Conf. Computer Vision* (ICCV), 2003, pp. 1-8.

[27] J. Sivic, and A. Zisserman Video Google: efficient visual search of videos, in *Toward Category-Level Object Recognition*, Springer LNCS, J. Ponce et al. (Eds.), Vol. 4170, pp. 127-144, 2006.

[28] J. Sun, M. Ovsjanikov, and L.J. Guibas, A concise and provably informative multi-scale signature based on heat diffusion, *Proc. Symposium on Geometry Processing*, 2009.

[29] R. Toldo, U. Castelllani, and A. Fusiello. The bag of words approach for retrieval and categorization of 3D objects. *The Visual Computer*, Vol. 26(10), Oct. 2010, pp. 1257-1268.

[30] T. Zaharia, and F. Preteux, 3D Shape-based retrieval within the MPEG-7 framework, *Proc. SPIE Conf. Nonlinear Image Proc. Pattern Anal.* XII, Vol. 4304, 2001, pp. 133-145.

[31] Y. Zhang and A.B. Hamza, Vertex-based diffusion for 3-D mesh denoising, *IEEE Trans. Image Processing*, Vol. 16(4), April 2007, pp. 1036-1045.

Chapter 6

Partial Shape Matching Using Graph Embedding and Its Application to Content-based Image Retrieval

6.1 Introduction

In spite of extensive research on shape matching over the years, obtaining a reliable and accurate correspondence between shapes in images that exhibit considerable variation in viewing conditions has proven to be an extremely difficult problem. In this chapter, we propose a multicriteria optimization-based technique for matching partially visible shapes in images with significant variation in viewing conditions. Classical approaches that attempt to solve the partial shape matching problem can be broadly categorized as global or local based on the granularity of the image features used. Most global shape matching techniques are seen to lack the ability to handle strong articulation, deformation or occlusion of objects (or their parts) whereas local feature-based matching are observed to fare poorly in the face of significant variations in illumination, scale, orientation and viewpoint. In this chapter, we focus

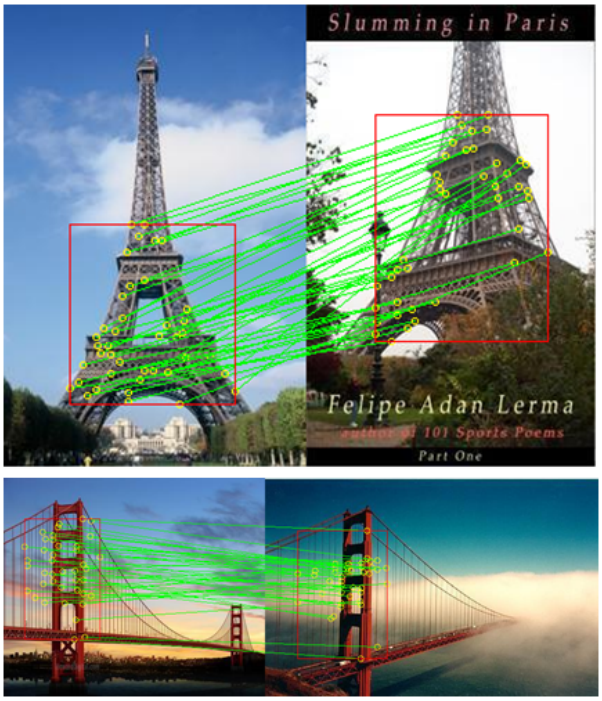

Figure 6.1: Partial shape matching over real world images. The result of the proposed technique is indicated by the red rectangle. Feature points are denoted using yellow circles whereas the feature point correspondences in the partial matches are shown using green lines.

on the problem of partial shape matching in images where the shapes of the objects exhibit significant variation on account of variations in viewing conditions. Given a set of features extracted from the images under consideration, our goal is to obtain accurate and optimal region-based correspondences for partial shape matching. Formal characterization of the partial shape matching problem is of particular importance because of its wide applicability in several computer vision problems such as structure-from-motion, object localization, fine-grained object categorization, shape-based image retrieval and image registration, to name a few.

Conventional global shape matching techniques compare two shapes by defining a global matching cost function which is then optimized to determine the best match. The shape

context technique proposed by Belongie et al. [9] has proven to be a popular method for global shape matching. It casts the shape matching problem as one of correspondence determination by computing a distance measure between the two shapes under consideration. However, the inability of global shape matching techniques to handle shape deformation arising from strong articulation or occlusion of objects (or their parts) has motivated the design of local shape matching techniques. Chen et al. [14] have addressed the shape deformation problem in their formulation of a local shape matching method but fail to provide a strong global description for accurate shape alignment. Ma and Latecki [30] have proposed a contour-based shape matching technique for real world images where the extent of shape articulation observed in the images is limited. Consequently, their technique is unable to handle the matching of highly articulated object shapes. Furthermore, their matching technique calls for prior knowledge of the underlying shape, severely limiting its scalability to real-world problems. Also, it must be noted that in most existing shape matching techniques, the underlying shape under consideration is often that of a very distinct object in an image captured in a highly controlled environment; an assumption that may not hold in many real world images.

In the proposed technique, we do not rely completely on either local shape matching methods based on local feature correspondences [14], [30] or on global shape matching techniques based on optimization of a global cost function [9]. Instead, we propose to use a multicriteria optimization technique for matching of persistent image features that encode the appearance similarity of objects or regions in a robust manner. As shown in this chapter, accurate partial shape matches in images that exhibit significant variations in imaging and viewing parameters can be obtained via optimization of a suitably defined distance measure between jointly embedded local feature vectors coupled with a strong global regularization constraint. In the proposed partial shape matching technique, we have investigated mean pixel intensity as the regularization criteria. The proposed technique is shown to not only

retrieve the area-wise partial similarity matching between the shapes, but also provide pixel-level correspondences between the shapes under consideration.

The primary motivation for the proposed framework comes from the recent work on the use of the eigenspectra of SIFT features [29] of the joint image graph as descriptors of image structure [5]. In particular, we propose the use of geometric blur (GB) features [11] as an alternative to SIFT features for generating the joint image graph. The inherent ability of GB features to focus on feature points on the dominant objects within the image has prompted their use, instead of SIFT features, for determining the partial geometric similarities between objects. The proposed partial shape matching technique has the potential to positively impact several complex computer vision applications. In particular we present the application of the proposed partial shape matching technique to a complex problem in computer vision, namely, content-based image retrieval (CBIR).

6.1.1 Content-based image retrieval (CBIR)

The field of CBIR is facing new challenges due to the rapid growth of image data available on the web, thanks to photo sharing sites such as Flickr and Picasa. Whereas traditional audiovisual archives come with carefully curated metadata, allowing easy access based on a predefined thesaurus, user-generated images are hardly annotated, except for a very few, and often with not very informative tags. The same holds for many older image archives that only recently have been digitized. This calls for methods that can analyze the images directly rather than relying on metadata.

In this chapter we formulate a CBIR scheme, based on the proposed partial shape matching technique, that can analyze the structure of the geometric shapes in the images and match them in a concise and meaningful manner directly, rather than relying on metadata. Conventional CBIR techniques are often confounded when the object of interest in the image is captured from different viewpoints or under different imaging conditions (e.g., different

lighting conditions).

Existing CBIR techniques can be broadly classified into two categories based on their degree of robustness to variations in viewpoint and imaging conditions. The first category is based on the frequency distribution of image features and commonly termed as *bag-of-visual-words* (BOVW)-based CBIR techniques [27]. BOVW-based techniques have been reported to yield good results under minimal variations in viewing conditions using the query-by-visual-example paradigm [49]. The second category of CBIR techniques focuses on invariant object and image descriptors to render the retrieval results robust to variations in imaging and viewing parameters. Affine-invariant representations [2], domain adaptation techniques [36], query expansion [16, 18] and multiple queries [21] have been proposed to enhance the robustness of CBIR. Even so, extreme variations in imaging and viewing conditions and cases of object deformation have been found to be a hindrance in the retrieval process.

In contrast to existing techniques, we pose CBIR as a problem in rich feature selection in order to enhance its accuracy and robustness. The proposed partial shape matching method is shown to be effective in tackling extreme variations in viewing conditions during the similarity-based retrieval process. The incorporation of the proposed partial shape matching method within a CBIR scheme entails the design of a training procedure based on selection of a rich feature set from *training image pairs* instead of a *single training image instance*. The training procedure adds two important characteristics to our CBIR framework. First, only features that are relevant to the object are selected; irrelevant features, such as those that describe the background, are ignored. Second, only features that are consistent across significantly different visual appearances and viewing conditions are selected. This results in a feature set that is both robust to variations in object appearance and viewing conditions and potentially most consistently descriptive of that particular object. The feature set thus obtained is further processed using a BOVW generation procedure followed by a discriminative binary coding (DBC) [37] procedure resulting in a compact binary code representation.

The Hamming distance between the binary codes of the query image and retrieved image is considered as the measure of (dis)similarity between the two and a k-nearest-neighbor (k-NN) search is used to determine the label of the query image.

6.1.2 Contributions of the chapter

The primary contributions of this chapter are as follows:

1. The formulation of a novel multicriteria optimization framework to address the inherently difficult problem of partial shape matching across images where the variations in imaging and viewing parameters are truly challenging.

2. The introduction of a new benchmark dataset wherein the variability ranges over several imaging and viewing parameters such as illumination (day versus night), viewpoint, age of structures (historic versus new), presence of occlusion and partially constructed structures, and inclusion of sketches and/or paintings of objects along with their captured images.

3. The application of the proposed partial shape matching framework to a challenging computer vision problem, i.e., content-based image retrieval.

6.2 Related Work

Shape matching techniques can be broadly categorized as global or local depending on the granularity of the features employed. Global shape matching techniques can be further categorized as contour-based and region-based methods. Contour-based methods exploit primarily the boundary information for matching the underlying shapes. One of the pioneering works in this area includes the *shape context* method proposed by Belongie et al. [9, 10]. For every sample point on the shape contour, the shape context method captures the spatial

distribution of all the other sample points relative to it. A distance measure is computed between the shapes under consideration and the shape matching problem is modeled as one of determining the optimal correspondence between the shapes. Shotton et al. [46] and Opelt et al. [35] use the chamfer distance measure to match fragments learned from codebooks to edge images. Zhu et al. [50] formulate the shape matching problem as a set matching problem. Ravishankar et al. [38] approximate the outer contours of objects by decomposing them into segments at points of high curvature which are then matched using a dynamic programming procedure. Bai et al. [4] use the shape skeleton to capture the primary structural information about the object followed by an oriented chamfer matching procedure [35] to match the object model parts to images. Other works in contour-based shape matching include the triangle area representation proposed by Alajlan et al. [1], and segment-based shape matching methods such as the shape tree method proposed by Falzenszwalb et al. [19] and the hierarchical procrustes matching procedure proposed by McNeill et al. [32].

Region-based approaches derive the shape descriptors using pixel-level information within a shape region. Some of the well known region-based shape matching methods are based on the computation of invariant moments such as the Zernike moments [25]. Other approaches such as the skeleton-based shape descriptors [42, 44] have proven better at capturing the articulation of shapes than conventional contour-based methods.

Although global shape matching methods, such as contour-based methods, are capable of capturing the global shape of the object, they are unable to handle strong articulations in the object shapes. Likewise, region-based methods, though reliable for complex shapes, rely primarily on global shape characteristics that exclude many important shape details. Consequently, these methods have less discriminative power when the intra-class feature variation is high. Skeleton-based global shape matching methods [42, 44] are able to handle limited shape articulation but, in the absence of region-based descriptors, their performance deteriorates significantly when confronted with complex shapes. As opposed to global shape

descriptors, local shape descriptors, such as the one proposed by Chen et al. [14], provide an accurate measure of local similarity, but fail to provide a strong global description for accurate shape alignment.

The proposed multicriteria optimization framework addresses primarily the problem of partial shape matching in real world images with extreme variations in viewing and imaging conditions. In addition to matching partially visible shapes, the proposed framework also improves the quality of feature point correspondences. Although several invariant feature-based matching techniques have been proposed for improving feature point correspondences, obtaining accurate, reliable and robust point correspondences is a challenging task using only local feature descriptors. One of the initial works in this area is the rotationally invariant descriptor proposed by Schmid and Mohr [41] which is constructed around a Harris corner detector using a set of Gaussian derivatives. Lowe [29] has extended this approach by incorporating scale invariance within the rotationally invariant descriptor.

In the context of determining feature point correspondences, Shi and Tomasi [45] have suggested that interest points should be located such that the solutions for the positions, orientations and scale of the matching feature points are stable. Patch matching techniques, such as the one proposed by Barnes et al. [6], perform well by determining the approximate nearest neighboring patches based on random sampling. However, extreme variations in viewing conditions encountered in real-world images render the patch matching technique infeasible. Other patch matching techniques include the multi-view matching method proposed by Brown et al. [13], wherein the matched oriented patches are observed to suffer from the same problem. A possible means to improve the feature correspondence is to incorporate mid-level cues such as the local shape of the region. Since the proposed multicriteria optimization framework optimizes the joint spectral embedding of features coupled with mean pixel intensity-based regularization, it restricts the features to a shape or region within the image resulting in improved correspondences between feature points.

Several methods for large-scale CBIR starting with a single query image have been proposed in the research literature [17, 31, 34]. These methods almost invariably comprise of the following components: bag-of-visual-words (BOVW) creation, indexing, query expansion and geometric verification. In query expansion (QE) [2, 16, 18, 21], the original query image is replaced with a more representative set of images that is constructed using the top-ranked retrieved images. This, in effect, turns the single-query retrieval problem into one of multiple-query retrieval. Geometric verification is typically used to improve the results by re-ranking the top-k retrieved images. The re-ranking step is linear in k, and therefore needs to be implemented efficiently, e.g. by using the LO-RANSAC procedure [17]. In addition, several improvements have been proposed in the research literature to the relatively standard CBIR components mentioned above. For instance, Zhang et al. [48] propose the use of geometry preserving visual phrases in their CBIR scheme. Although their method captures geometric information at the level of feature construction level, it is sensitive to local deformations as the structures it discovers are inherently rigid.

6.3 Theoretical Derivation

6.3.1 Partial Shape Matching

We first define the class of objects we are interested in comparing across image pairs. The proposed technique focuses on image pairs containing a *single dominant object* where the dominant object is either completely or partially visible. We denote the images constituting the image pair under consideration as X and Y. We assume that the images X and Y can each be decomposed into subparts, where the subparts are modeled as subsets $X' \subset X$ and $Y' \subset Y$ respectively. We establish the *equivalence relation* (denoted by $\sim$) between the two subparts X' and Y' (instead of *exact similarity* which implies $X' = Y'$) under certain conditions. From now on, with a little abuse of notation, we will use the terms *similarity*

and *equivalence* interchangeably.

The *degree of dissimilarity* between subparts X' and Y' can be expressed by a non-negative function $d : \Sigma'_X \times \Sigma'_Y \rightarrow \mathbb{R}+$, where Σ defines the geometry of the subpart. We design the first term of the objective function in the proposed multicriteria optimization framework as a dissimilarity measure between feature pixels at a local geometric level. The second term in the objective function is a regularization term comprising of an intensity-based similarity measure which encapsulates the global structure of the shape. For a feature point $x \in X$, region-wise matching can be achieved using the proposed multicriteria optimization framework as follows:

$$\min_{X',Y';\varphi:X'\rightarrow Y'} \left[\left\{ \int_{X'\times Y'} d(x, \varphi(x)) \right\} + r(X', Y') \right] \tag{6.1}$$

Joint Geometric Embedding

The spectral analysis of the contents of an image is typically performed on a weighted image graph $G(V, E, W)$ [3]. The vertices of the graph V denote the pixel-level features of the image. The edge set E denotes the pair-wise relationships between every pair of vertices in the set V, making G a complete graph. The weight $w_{ij} \geq 0$ associated with an edge $(v_i, v_j) \in E$ encodes the affinity between the corresponding pixel-level features represented by vertices v_i and v_j. The edge weights are represented by an $n \times n$ affinity matrix $W = [w_{ij}]_{i,j=1,2,\ldots,n}$.

The above formulation is extended for a joint graph as follows: Let $G_1(V_1, E_1, W_1)$ and $G_2(V_2, E_2, W_2)$ be the image graphs for images I_1 and I_2, respectively. The joint image graph $G(V, E, W)$ is defined such that $V = V_1 \cup V_2$ and $E = E_1 \cup E_2 \cup (V_1 \times V_2)$ where $V_1 \times V_2$ is the set of edges connecting every pair of vertices in (V_1, V_2). The resulting affinity matrix W is given by:

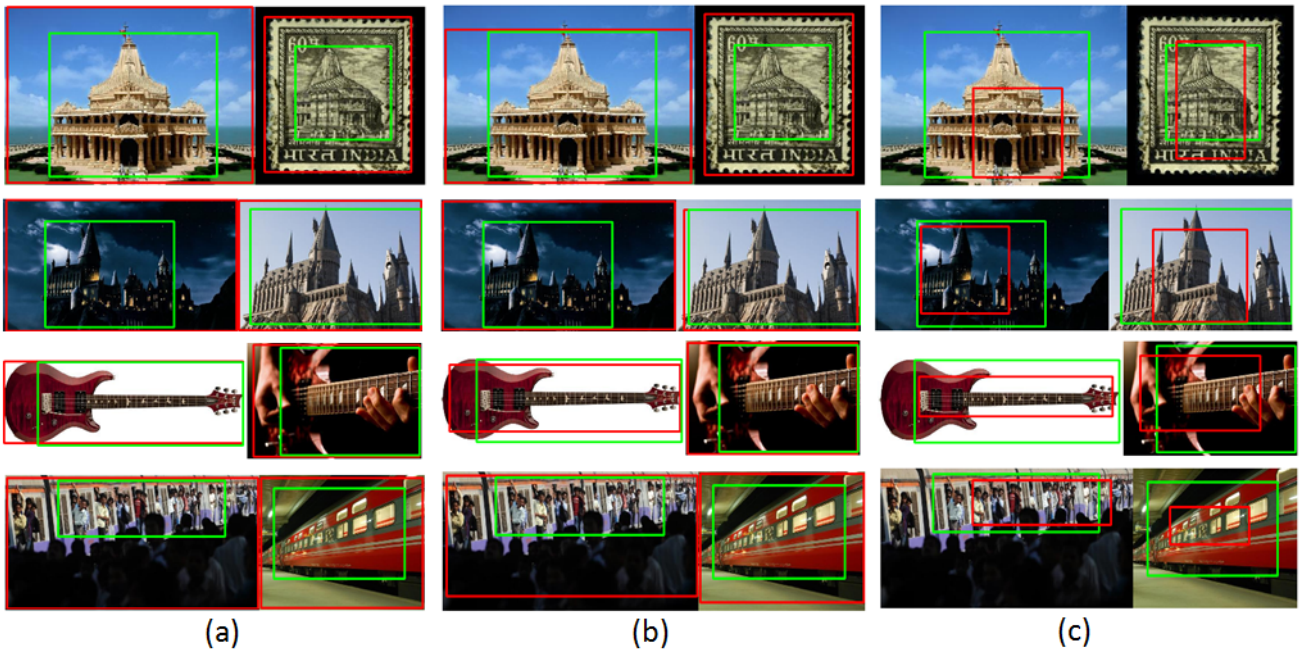

Figure 6.2: Partial shape matching in real world images under varying conditions. The green rectangle shows the ground truth partial match whereas the results of using (a) SIFT features, (b) GB features and (c) the proposed method are denoted using the red rectangle.

$$W = \begin{bmatrix} W_1 & C \\ C^T & W_2 \end{bmatrix}_{(n_1+n_2)\times(n_1+n_2)} \tag{6.2}$$

where the affinity sub-matrices W_1, W_2 and C are defined as follows:

$$(W_i)_{x,y} = \exp(-(\|f_i(x) - f_i(y)\|)^2 \tag{6.3}$$

$$C_{x,y} = \exp(-(\|f_i(x) - f_j(y)\|)^2 \tag{6.4}$$

where, $f_i(x)$ and $f_i(y)$ are pixel-level features at locations x and y respectively in image I_i and $\|.\|$ is the Euclidean norm or distance measure.

We consider geometric blur (GB) [11] as the feature descriptor of choice since our primary objective is to focus on the geometric structure of the dominant object within the image. GB features yield image descriptors that are robust to small transformations. GB averages the underlying signal over small transformations and then samples the signal at fixed locations to construct the desired robust descriptor. For oriented images, the edge response has been shown to be a good underlying signal. In practice, the averaging of transformations can be modeled by convolving the signal with a kernel that weighs the contribution of neighboring signals at a given point or pixel location. Gaussian kernels where the support (σ) increases linearly with the distance from the origin are commonly used for this purpose [11].

We compute the *joint geometric embedding distance* (JGED) between two feature points x and y in two different images (comprising the image pair) using the first m non-trivial eigenvectors (ϕ_k) of the joint graph as follows:

$$d^2_{JGED}(x, y) = \sum_{k=1}^{m} (\phi_k(x) - \phi_k(y))^2 \tag{6.5}$$

Computation of d_{JGED} in the embedding space ensures that the resulting distances are more robust to the variations in imaging and viewing parameters commonly encountered in real-world images than the feature distances computed in the GB space. This joint geometric embedding distance is the first term of the proposed multicriteria optimization framework (eq. (6.1)).

6.3.2 Regularization

The quality of the match between subparts X' and Y' in the image pair under consideration can be measured using an appropriately defined region-based *irregularity function* $r(X', Y')$. Minimization of the region-based irregularity term ensures a sufficiently high-quality match between the two regions instead of just the feature points, i.e., $r(X', Y')$ serves as a regu-

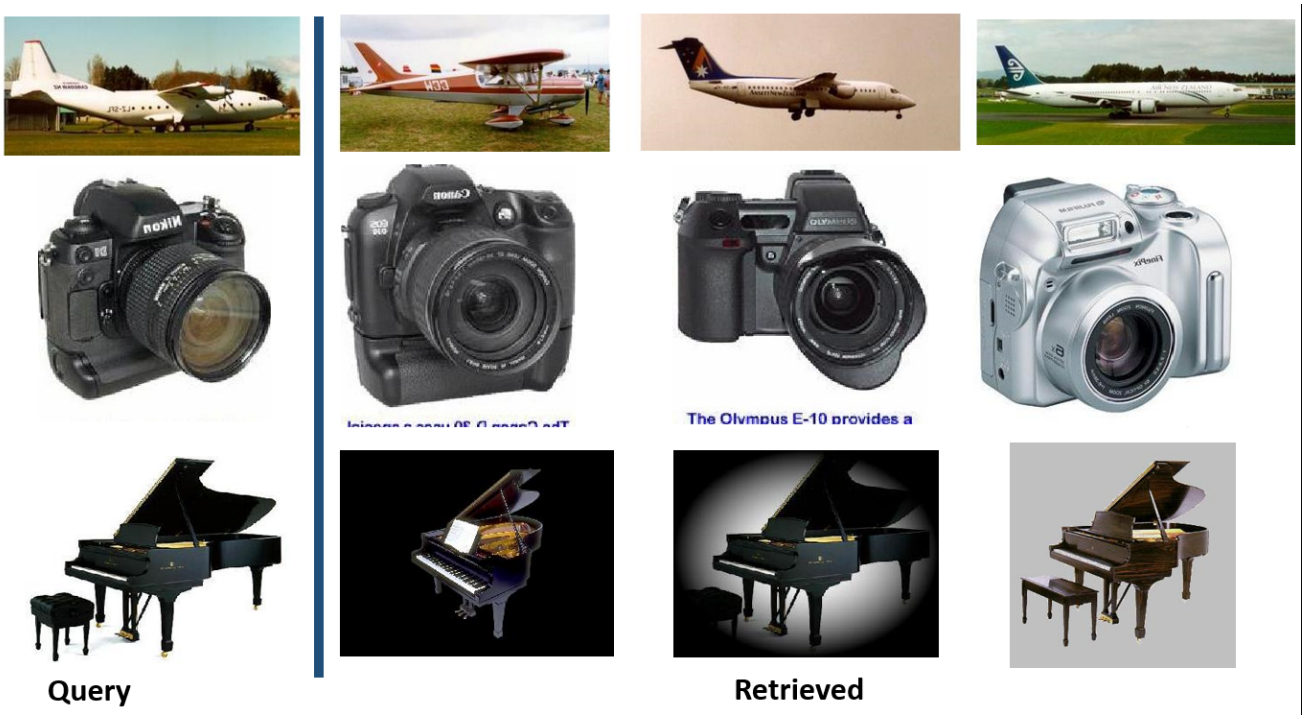

Figure 6.3: The differences in object appearances in a CBIR system.

larization term in the multicriteria optimization framework (eq. (6.1)). In particular, we compute the difference in values of mean pixel intensity of the regions X' and Y' to regularize the quality of match between the corresponding subparts in the image pair.

6.4 Content-based Image Retrieval

Content-based image retrieval (CBIR) techniques directly analyze the contents of a query image and return images with similar content from an image database. A typical CBIR system comprises of (a) efficient retrieval of images from the image database/archive that are similar to a query image using an inverted file system or an index structure based on an R/R*-tree [8, 24] and (b) generation of a ranked list of the retrieved images, with images ranked based on their similarity to the given query image. However, since an object in

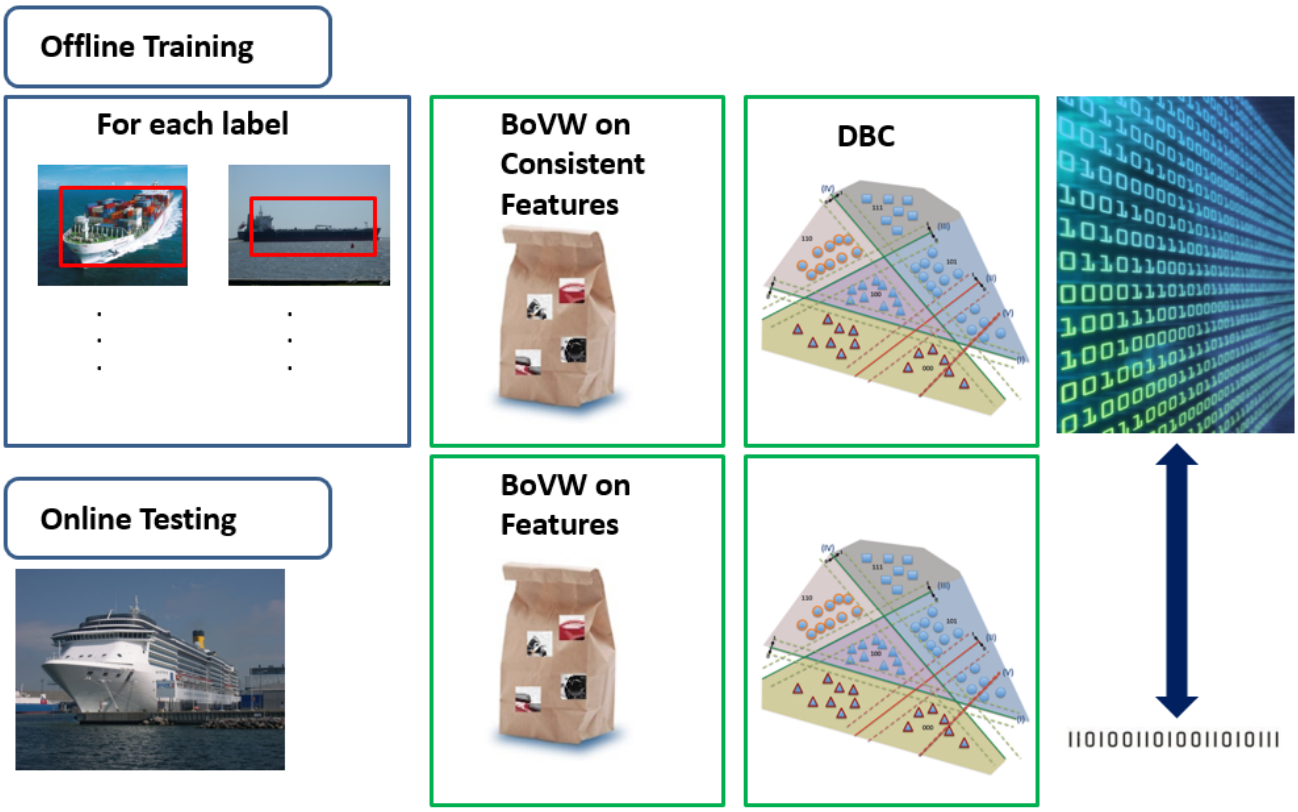

Figure 6.4: An outline of the proposed procedure for Content-based Image Retrieval.

an image may be captured from different viewpoints or under different viewing conditions (Figure 6.3), the quality of retrieval is very sensitive to the query image. The dependence of the retrieval quality on the query image has been alleviated by developing more robust image representations that are invariant to changes in viewing conditions. Affine invariant representations [2], domain adaptation techniques [36], and techniques such as query expansion [16, 18] and multiple queries [21] have been used with reasonable success. However, extreme changes in viewing conditions still pose a significant challenge to most CBIR systems.

The proposed partial shape matching technique, on the other hand, is shown to be capable of detecting similar shapes between images that vary significantly in terms of illumination,

viewpoint, age of structures, presence of occlusion and partially constructed structures, and inclusion of sketches and/or paintings of objects along with their captured images. It is a fair observation that the proposed partial shape matching framework is able to find stable geometric features across images in the face of highly varying imaging and viewing conditions. The proposed CBIR framework is designed to leverage the proposed partial shape matching technique to discover and extract consistent features across extreme variations in imaging and viewing conditions to ensure more robust and accurate representation and retrieval of images from the image archive/database containing the object of interest captured in the query image.

6.4.1 Offline Training

The robust and consistent features for representation of an object in an image are learned using an offline training procedure on the archived images. For each class of images containing a single specific dominant object of interest, a set of image pairs is randomly generated. For each pair of images, the proposed partial shape matching procedure is performed to localize the dominant object in the images and establish the feature point correspondences between the object instances. The feature point correspondences thus obtained are used to determine the consistent features across the varying imaging and viewing conditions represented by the image pair. The extraction of consistent features from image pairs that exhibit good shape correspondence for a specific dominant object, enables the generation of a rich feature set for shape description that could used for the purpose of retrieval. The feature extraction phase is followed by a bag-of-visual-words (BOVW) technique to generate the frequency histograms of the consistent spatial features. These consistent features for each image class are learned during the offline training process resulting in a feature-level description for the corresponding object class, as depicted in Figure 6.4.

Discriminative Binary Code Generation

For representation and rapid matching of features during online retrieval, the discriminative binary code (DBC) developed by Rastegari et al. [37] is employed. The main purpose of the DBC is to represent features in a reduced binary space which retains some similarity to the original real-valued feature space. The goal here is to learn the binary codes for each instance in the training set such that:

1. The resulting DBCs can be reliably predicted from the visual data and,

2. Representation of each image with its learned DBC enables rapid and accurate discrimination of the image.

The proposed CBIR system comprises of two stages similar to the ones described in [37]: (a) learning the DBC for each feature instance and, (b) performing search or classification in the space of DBCs. The DBC learning procedure is based on the joint optimization of two criteria outlined in [37]: the resulting binary code should be (a) maximally discriminative and, (b) maximally predictive of the visual data (i.e., the underlying features). The above two criteria impose conflicting constraints, for example, assigning a unique code to examples within the same category would result in the most discriminative codes but ones that would be extremely hard to predict from visual data. The most predictable codes on the other hand, may contain very little information about the underlying categories resulting in poor discrimination. The DBC generation procedure in [37] provides a trade-off between the discriminative and predictive capabilities of the resulting codes.

Following [37], a code is deemed discriminative if the coded instances of different categories are distant from one another whereas those belonging to the same category are proximal based on some distance measure. However, these discriminative constraints are not encoded as hard constraints; rather a DBC is assigned to each feature instance or image in a way that the resulting DBCs have enough discriminative power and yet can be reliably

predicted from images. Such a code allows for simple, efficient and accurate classification using a Support Vector Machine (SVM) and retrieval using k-NN search technique. Each bit of the DBC can be visualized as a hyperplane that separates feature instances or images that have a value 0 versus the ones that have a value 1. Each bit of the DBC is generated by checking which side of a hyperplane a feature instance or image lies on [37].

6.4.2 Online Testing

During the testing phase, each query image is processed using the GB feature extraction and BOVW generation procedures. The BOVW representation is further processed using the DBC generation procedure resulting in a DBC representation for each query image. During the retrieval process, a k-NN search is performed using the Hamming distance between the DBC representation of the query image and the DBC representations of the images in the image database/archive and the mode of k-NN is considered as the retrieval result.

6.5 Experimental Results

We designed a set of experiments to evaluate the proposed partial shape matching technique. We evaluated the performance of the proposed technique using two different metrics. The first metric evaluates the repeatability of the proposed technique for different image pairs whereas the second metric compares the reliability of the proposed technique with that of scale invariant feature transform (SIFT) [11], geometric blur (GB) [11], speeded up robust features (SURF) [7], features from accelerated segment test (FAST) [39] and oriented FAST and oriented BRIEF (ORB) [40] features on a partial shape matching benchmark designed by us. Since certain correspondence determination techniques [30] require the contour model to be known a priori, it was not feasible to evaluate and compare the results of these techniques using our benchmark dataset since it includes wide variations in viewing conditions.

For the purpose of evaluation, we collected a set of image pairs with a *single dominant object* that is either completely or partially visible. These images focus on architectural scenes and, in some cases, musical instruments, which we believe will prove extremely challenging for current feature matching methods. The image pairs exhibit an array of dramatic variations in object appearances, arising due to variations in illumination, age of the objects and rendering styles (paintings, drawings, sketches etc.), occlusion, presence of structures still under construction, to name a few. None of the image pairs are pre-aligned, thus ensuring that factors such as geometry and appearance variations play an important role in the matching process. In total, we gathered 40 image pairs (`http://cobweb.cs.uga.edu/~csc/Supplimentary.rar`), all of which are the result of Google Image Search. For each image pair we manually annotated a bounding box around the dominant object (denoted by X'_g and Y'_g) in each of the images for use in computing the ground truth matches. Examples of ground truth matches between the image pairs are shown in Figure 6.1 and Figure 6.2.

6.5.1 Evaluation of the Matching Procedure

For objective evaluation of the performance of the proposed technique, we computed a *relevance score* R for each image pair as follows:

$$R = \frac{TP}{TP + TN + FP} \tag{6.6}$$

where,

$$TP = (X' \cap X'_g) \cup (Y' \cap Y'_g) \tag{6.7}$$

$$TN = (X'_g - (X' \cap X'_g)) \cup (Y'_g - (Y' \cap Y'_g)) \tag{6.8}$$

$$FP = (X' - (X' \cap X'_g)) \cup (Y' - (Y' \cap Y'_g)) \tag{6.9}$$

The mean relevance score (MRS) and the false positive rate (FPR) values were considered as the metric for performance evaluation of all image pairs in the benchmark.

6.5.2 Parameter Selection

The two main parameters that need to be chosen for the proposed geometric embedding method are n_{GB} (number of GB feature points) and ϕ (the number of top eigenvectors). We computed the MRS and FPR values over all the image pairs for different combinations of values for the parameters n_1 and ϕ. The results are reported in Table 6.1.

Since the combination of parameter values $n_{GB} = 300$, $\phi = 20$ and mean intensity-based regularization was observed to yield the best classification accuracy, we used this set of parameter values for all the experiments reported in the remainder of the chapter unless mentioned otherwise.

6.5.3 Repeatability of the Matching Procedure

The random initialization step in the procedure for extraction of the geometric blur feature points necessitates an evaluation of repeatability of the proposed partial shape matching technique on our benchmark dataset. For each image pair, a *repeatability score* R_1 was computed as follows:

$$R_1 = \frac{(X' \cap X'_g) \cup (Y' \cap Y'_g)}{X' \cup Y'} \tag{6.10}$$

We ran the proposed partial shape matching procedure 50 times for each image pair. The value of the repeatability score R_1 was computed for each run for each image pair and plotted in Fig. 6.5. It is important to note that in the case of the first outlier, i.e., the *Empire*

Table 6.1: Parameter tuning for maximizing the MSR and minimizing the FPR across 40 image pairs with mean intensity-based regularization.

n_{GB}	ϕ	MRS	FPR
100	5	0.71	0.27
100	10	0.72	0.26
100	20	0.76	0.22
100	30	0.76	0.22
200	5	0.75	0.23
200	10	0.77	0.21
200	20	0.79	0.19
200	30	0.8	0.18
300	5	0.74	0.24
300	10	0.78	0.2
300	20	0.82	0.17
300	30	0.81	0.18

State Building, the scarcity of prominent features in the left image resulted in a very low repeatability score; whereas in the case of the second outlier, i.e., the *Sphinx*, the GB features considered the people occluding the Sphinx in the right image to constitute the dominant object, resulting in low repeatability score. The average repeatability score over the entire dataset (excluding the two outliers mentioned above) was observed to be 0.8255 which is quite high and shows the effectiveness of the proposed partial shape matching procedure.

6.5.4 Comparison

We evaluated the results of the proposed partial matching technique and compared them with those obtained using standard features like SIFT [29], geometric blur (GB) [11], speeded up robust features (SURF) [7], features from accelerated segment test (FAST) [39] and oriented FAST and oriented BRIEF (ORB) [40] in isolation. In the case of most image pairs, using the aforementioned standard features in isolation failed to localize the partial shape of the

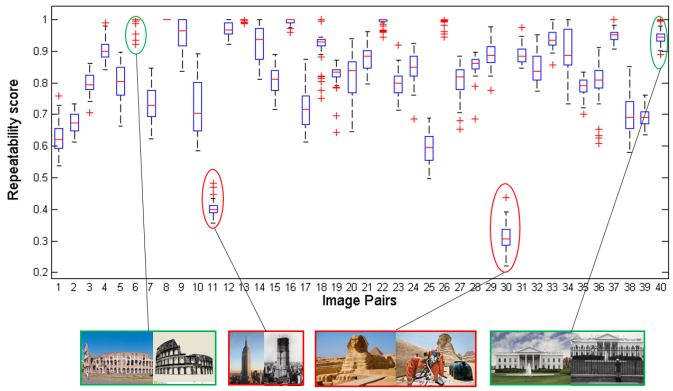

Figure 6.5: Repeatability score R_1 for all the image pairs after 50 runs per image pair. Two instances of image pairs with high repeatability scores (green) and two instances of outliers (red) are also shown.

Table 6.2: Comparison of the mean relevance score (MRS - higher is better) and false positive rate (FPR - lower is better) for standard feature-based techniques and the proposed technique.

	SIFT [29]	GB [11]	ORB [40]	SURF [7]	FAST [39]	Proposed
MRS	0.49	0.36	0.63	0.59	0.40	0.82
FPR	0.48	0.61	0.34	0.37	0.54	0.17

object. The quantitative evaluation is summarized in Table 6.2. It is particularly important to note that the proposed technique, on average, performed significantly better than the techniques that use standard feature-based techniques in isolation. In almost all cases, the standard feature-based matching techniques were observed to regard the entire image as a partial match, thus rendering them ineffective in solving the partial shape matching problem.

6.5.5 Content-based Image Retrieval Results

Datasets

We tested the proposed CBIR method on the Caltech256 dataset [22]. This dataset has a large number of categories (256) containing significant intra-class variations. Category

retrieval in the case of the Caltech256 dataset is a challenging task since the number of categories is much higher than typical number of experiments and also the intra-class variations are much higher than those encountered in typical datasets. There are around 30,000 images in the Caltech256 dataset spread over 256 categories with 120 images per category on average.

Evaluation Procedure

To evaluate the proposed CBIR technique we performed a series of extensive evaluations and comparisons. For our method, we varied the following settings: the length of binary codes $b \in 32, 64, 128, 256, 512$ and the number of training examples per category in the range $[5, 50]$. In the interest with methods in the literature we compared the mean retrieval precision of the proposed method with that of the Locality Sensitive Hashing (LSH) technique as a standard baseline, with the supervised version of Iterative Quantization (ITQ) [23] as the best supervised method and Spectral Hashing (SpH) [47] as the state-of-the-art unsupervised method for producing binary codes. Varying $k = [1 : 5 : 100]$ for the k-NN search we computed a set of precision values. The mean precision value was obtained by taking the average of these precision values. The retrieval performance was measured by plotting the mean retrieval precision versus the number of training images. Our experimental evaluations in Figure 6.6 demonstrate that our method consistently outperforms state-of-the-art methods under all the combinations of above settings.

6.6 Conclusions and Future Work

In this chapter, we presented a novel multicriteria optimization framework for partial shape matching in general real-world images. The proposed framework makes two primary contributions: First, a multicriteria optimization technique for partial shape matching suitable for

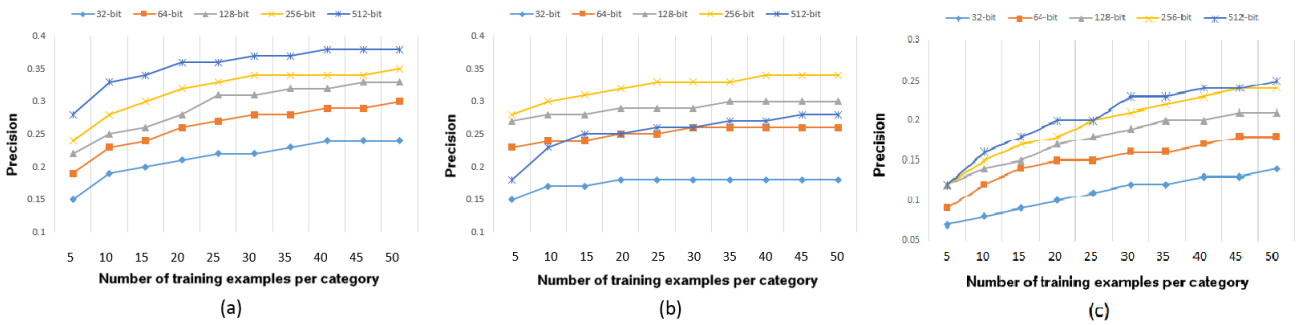

Figure 6.6: Our method produces state-of-the-art results on the Caltech256 dataset. This figure compares the performance in terms of retrieval precision versus the number of training examples per category on the Caltech256 dataset of (a) the proposed technique with (b) the supervised version of Iterative Quantization (ITQ) [23] and (c) Spectral Hashing (SpH) [47].

matching images under dramatically different imaging and viewing conditions is proposed and successfully applied to a difficult problem in computer vision, i.e. Content-based Image Retrieval (CBIR). Second, a novel benchmark is presented here, containing image pairs of ground truth partial matches where the underlying variability includes several factors such as illumination, viewpoint, age of structures, occlusion, structures under partial construction and inclusion of paintings and sketches of objects. Performing the region-wise optimization rather than finding point correspondences is the main reason behind the success of our algorithm and the results on state-of-the art benchmark datasets for the CBIR application bolster our claims. In future, we plan to extend the proposed framework by incorporating 2.5D data into the framework and considering more complex computer vision problems such as structure-from-motion.

Bibliography

[1] N. Alajlan, I. E. Rube, M. S. Kamel, and G. Freeman, Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition* 40, pp. 19111920, 2007.

[2] R. Arandjelovic and A. Zisserman. Three things everyone should know to improve object retrieval. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 2911- 2918, 2012.

[3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, Contour detection and hierarchical image segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 33(5), pp. 898 - 916, 2011.

[4] X. Bai, L. J. Latecki, and W. Y. Liu. Skeleton pruning by contour partitioning with discrete curve evolution. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 29(3): 449-462, 2007.

[5] M. Bansal and K. Daniilidis, Joint spectral correspondence for disparate image matching, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 2802-2809, 2013.

[6] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman, PatchMatch: a randomized correspondence algorithm for structural image editing, *ACM Trans. Graphics* 28(3):24, 2009.

[7] H. Bay, T. Tuytelaars, and L. Van Gool, Surf: Speeded up robust features. *Proc. European Conf. Computer Vision (ECCV)*, pp. 404-417, 2006.

[8] N. Beckmann, H. P. Kriegel, R. Schneider and B. Seeger, The R*-tree: an efficient and robust access method for points and rectangles, *Proc. ACM SIGMOD*, pp. 322-331, 1990.

[9] S. Belongie, J. Malik, and J. Puzicha, Shape context: A new descriptor for shape matching and object recognition, *Advances in Neural Information Processing Systems (NIPS)*, Vol. 2, p. 3, 2000.

[10] S. Belongie, J. Malik, and J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 24(4), pp. 509-522, 2002.

[11] A. Berg and J. Malik, Geometric blur for template matching, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Vol. 1, pp. 607-614, 2001.

[12] A. M. Bronstein, M. M. Bronstein, A. M. Bruckstein, and R. Kimmel, Partial similarity of objects, or how to compare a centaur to a horse, *Intl. Jour. Computer Vision*, Vol. 84(2), pp. 163-183, 2008.

[13] M. Brown, R. Szeliski, and S. Winder, Multi-image matching using multi-scale oriented patches, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 510-517, 2005.

[14] L. Chen, R. Feris and M. Turk, Efficient partial shape matching using Smith-Waterman algorithm, *Proc. NORDIA Workshop at CVPR*, pp. 1-6, 2008.

[15] O. Chum, J. Matas, and S. Obdrzalek. Enhancing RANSAC by generalized model optimization. *Proc. Asian Conf. Computer Vision (ACCV)*, pp. 812-817, 2004.

[16] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman. Total recall: Automatic query expansion with a generative feature model for object retrieval. *Proc. International Conf. Computer Vision (ICCV)*, pp. 1-8, 2007.

[17] O. Chum, J. Philbin, and A. Zisserman. Near duplicate image detection: min-hash and tf-idf weighting. *Proc. British Machine Vision Conf. (BMVC)*, pp. 812-815, 2008.

[18] O. Chum, A. Mikulik, M. Perdoch, and J.Matas. Total recall ii: Query expansion revisited. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 889-896, 2011.

[19] P. F. Felzenszwalb and J. Schwartz, Hierarchical matching of deformable shapes. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 1-8, 2007.

[20] P. F. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, Object detection with discriminatively trained part based models, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 32(9), pp. 1627 1645, 2010.

[21] B. Fernando and T. Tuytelaars. Mining Multiple Queries for Image Retrieval: On-the-fly learning of an Object-specific Mid-level Representation. *Proc. International Conf. Computer Vision (ICCV)*, pp. 2544-2551, 2013.

[22] G. Griffin, A. Holub, and P. Perona: The Caltech-256. Technical report, 2007.

[23] Y. Gong, S. Lazebnik, Iterative quantization: A procrustean approach to learning binary codes. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 817-824, 2011

[24] A. Guttman, R-Trees: A dynamic index structure for spatial searching, *Proc. ACM SIGMOD*, pp. 47-57, 1984.

[25] W. Y. Kim, Y. S. Kim, A region-based shape descriptor using Zernike moments. *Signal Process.: Image Commun.* 16 (1), 95-102, 2000.

[26] A. Kuijper and O. F. Olsen, Describing and matching 2D shapes by their points of mutual symmetry, *Proc. European Conf. Computer Vision (ECCV)*, pp. 213-225, 2006.

[27] F-F. Li and P. Perona, A Bayesian hierarchical model for learning natural scene categories, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)* Vol. 2, pp. 524-531, 2005.

[28] Z. Liu, H. Shena, G. Fengb, and D. Hua, Tracking objects using shape context matching, *Neurocomputing* pp. 47-55, 2012.

[29] D. G. Lowe, Distinctive image features from scale-invariant keypoints *Intl. Jour. Computer Vision (IJCV)*, Vol. 60(2), pp. 91-110, 2004.

[30] T. Ma and L. J. Latecki, From partial shape matching through local deformation to robust global shape similarity for object detection, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 1441-1448, 2011.

[31] A. Makadia. Feature tracking for wide-baseline image retrieval. *Proc. European Conf. Computer Vision (ECCV)*, pp. 310-323, 2010.

[32] G. McNeill, S. Vijayakumar, Hierarchical procrustes matching for shape retrieval. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 885-894, 2006.

[33] G. Mori, S. Belongie, and J. Malik, Shape contexts enable efficient retrieval of similar shapes, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 723-730, 2001.

[34] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 2161-2168, 2006.

[35] A. Opelt, A. Pinz, and A. Zisserman, A boundary-fragment-model for object detection, *Proc. European Conf. Computer Vision (ECCV)*, pp. 575-588, 2006.

[36] J. C. Pereira and N. Vasconcelos, Cross-modal domain adaptation for text-based regularization of image semantics in image retrieval systems, *Computer Vision and Image Understanding (CVIU)*, Vol. 124, pp. 123135, 2014.

[37] M. Rastegari, A. Farhadi, and D. Forsyth, Attribute discovery via predictable discriminative binary codes, *Proc. European Conf. Computer Vision (ECCV)*, pp. 876-889, 2012.

[38] S. Ravishankar, A. Jain, and A. Mittal. Multi-stage contour based detection of deformable objects. *Proc. European Conf. Computer Vision (ECCV)*, pp. 483-496, 2008.

[39] E. Rosten, and T. Drummond, Machine learning for high-speed corner detection, *Proc. European Conf. Computer Vision (ECCV)*, pp. 430-443, 2006.

[40] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, ORB: an efficient alternative to SIFT or SURF, *Proc. Int. Conf. Computer Vision (ICCV)*, pp. 2564-2571, 2011.

[41] C. Schmid and R. Mohr, Local grayvalue invariants for image retrieval, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 19(5), pp. 530-535, 1997.

[42] T.B. Sebastian and B.B. Kimia, Curves vs skeletons in object recognition, *Proc. Intl Conf. on Image Processing (ICIP)*, pp. 247-263, 2001.

[43] T. Sebastian and P. N. Klein, On aligning curves, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 25(1), pp. 116-125, 2003.

[44] T. Sebastian, P. Klein, and B. Kimia, Recognition of shapes by editing their shock graphs, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 26(5), pp. 550-571, 2004.

[45] J. Shi and C. Tomasi, Good features to track, *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp.593-600, 1994.

[46] J. Shotton, A. Blake, and R. Cipolla. Multi-scale categorical object recognition using contour fragments. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, pp. 503-510, 2008.

[47] Y. Weiss, A.B. Torralba, R. Fergus, Spectral hashing. *Advances in Neural Information Processing Systems (NIPS)*, pp.1753-1760, 2008.

[48] Y. Zhang, Z. Jia, and T. Chen. Image retrieval with geometrypreserving visual phrases. *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 809-816, 2011.

[49] J. Zhang, M. Marszaek, S. Lazebnik, C. Schmid, Local features and kernels for classification of texture and object categories: a comprehensive study, *Intl. Jour. Computer Vision*, Vol. 73(2), pp. 213238, 2007.

[50] Q. Zhu, Y. W. L. Wang, and J. Shi. Contour context selection for object detection: A set-to-set contour matching approach. *Proc. European Conf. Computer Vision (ECCV)*, pp. 774-787, 2008.

Chapter 7

Discussion and Conclusions

In this final chapter, we first summarize what the dissertation has achieved thus far. The second section highlights the general contribution of the dissertation towards the area of computer vision and graphics as well as the interdisciplinary area of medical image analysis. In the third and final section, the scope and directions for future research are presented. It is important to note that, as each chapter contains its individual contributions and future directions, in this chapter we only state the major future research directions.

7.1 Summary

The primary goal of the dissertation was to understand shapes from a geometric perspective and determine computational geometric invariants for the deformations the shape has undergone. Moreover, the focus of the dissertation was to develop mid-level and high-level understanding of shape analysis to enable direct applicability of the theoretical insights to practical applications. As a result, this dissertation provided some important theoretical insights in to the field of shape analysis as well as some initial directions towards applying these insights to address some real world applications.

In particular, the four main theoretical developments presented in this dissertation are Geodesic Field Estimate, Biharmonic Density Estimate, Intrinsic Symmetry Detection and Partial Shape Matching. Moreover, we have proposed a 3D Bag-of-Features (BoF) model for quantifying the endocardial structure of the left ventricle, a Geodesic Context Histogram-based descriptor based on the GFE for analyzing cortical folding patterns in the human brain and an efficient Content-based Image Retrieval technique using partial shape matching to enhance the applicability of theoretical knowledge to practical problems.

Chapter 2 introduced the Geodesic Field Estimate (GFE), a GPU-optimized version of all pairs shortest path determination algorithm and the Geodesic Context Descriptor (based on GFE) which in turn is used to quantify the regularity of the cortical surface of the human brain. Biharmonic Density Estimate, another novel low-level descriptor was introduced in Chapter 3. In Chapter 4, we developed a framework for detecting and characterizing intrinsic symmetry over 3D shapes. Chapter 5 introduced a 3D BoF technique for quantifying the left ventricular endocardial morphology whereas Chapter 6 developed a partial shape matching technique which in turn is applied to the Content Based Image Retrieval task.

7.2 General Contributions

As mentioned earlier, this dissertation provides a deep theoretical and practical understanding of shape analysis in particular and computer vision and graphics in general. The general theoretical contributions which we believe can have a lasting impact are listed below:

1. Introduction of the GFE, an intrinsic geometric surface signature based on the GFE, termed as the GFE contextual histogram (GCH), as well as the design and implementation of a parallel version of the all pairs geodesic path determination algorithm, a key element in the GFE computation, using GPUs to ensure that the GFE computation is indeed scalable for large datasets.

2. Introduction of the BDE that results in the formulation of a multiscale 3D shape signature as a distance density function, that can effectively capture both, global shape and local surface detail, depending on the scale parameter value. Moreover, the BDE is one of the first works where the intrinsic geometric scale space parameter can be directly related to the neighborhood size r which is both, intuitive and easy to relate to the perceivable 3D surface geometry.

3. Robust and meaningful characterization of the symmetry transformation via formulation of a symmetry space which allows one to quantitatively distinguish between instances of simple and complex intrinsic symmetry. To the best of our knowledge, this quantitative distinction has not been attempted in the published literature.

4. The formulation of a novel multicriteria optimization framework to address the hitherto unsolved problem of partial shape matching across images where the variations in imaging and viewing parameters are truly challenging.

5. The adaptation of the 3D Bag-of-Features (BoF) framework for non-rigid shape analysis for the purpose of cardiac shape analysis which is an important problem in cardiac imaging in particular and medical imaging in general.

6. The formulation of the partial shape matching framework for solving the challenging computer vision problem of content-based image retrieval.

7. The characterization of geometric regularity and variability of the highly irregular cortical surface folding patterns in the human brain at the DICCCOL sites using GCH.

7.3 Future Research

We have successfully developed two intrinsic geometric features, which can successfully describe a shape and its deformations. But, considering that the BDE is solely dependent on the modal analysis of the Laplace-Beltrami operator, one possible future theoretical direction can be to focus on modal analysis of other operators which have the potential to be invariant under a more complex deformations than just intrinsic isometry.

From the intrinsic symmetry detection perspective, since the proposed technique is equipped with functional maps, in future it will be possible to formulate operations, such as addition and subtraction, on the generated functional maps that could potentially provide a deeper and more comprehensive understanding of intrinsic symmetry in general. One major challenge towards quantifying the endocardial surface structure using the 3D BoF is that, it is not possible to substantially localize certain endocardial surface changes. Since the BoF histogram practically removes all the spatial information. As a result, it would be clinically more meaningful to devise methods in future that can successfully localize problem areas on the LV endocardial suraface.

Finally, the partial shape matching framework has potential applications in many high-level computer vision problems. We have only applied the technique to deal with Content Based Image Retrieval thus far. But, in future, we plan to extend this method towards solving problems such as Fine-Grained Recognition, Structure from Motion and wide-baseline 2D-2D or 3D-2D registration of urban scenes.