

LBO-Shape densities: A unified framework for 2D and 3D shape classification on the hypersphere of wavelet densities



Mark Moyou^{a,*}, Koffi Eddy Ihou^b, Adrian M. Peter^a

^a Dept. of Engineering Systems, Florida Institute of Technology, Melbourne, FL, USA

^b Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Qc, Canada

ARTICLE INFO

Article history:

Received 10 September 2015

Revised 29 April 2016

Accepted 27 July 2016

Available online 9 August 2016

Keywords:

2D and 3D shape matching and retrieval

Laplace–Beltrami Operator

square-root wavelet density estimation

Graph Laplacian

Intrinsic shape representation

Karcher mean

ABSTRACT

Recent years have seen a sharp rise in shape classification applications, and following suit, several frameworks have been proposed for efficient indexing of shape models. Here we propose a state-of-the-art shape matching framework which concomitantly provides transformation invariance and computationally efficient querying. Shapes are represented as probability density functions estimated on the eigenspace of the shape's Laplace–Beltrami operator (LBO) and the ensuing manifold geometry is leveraged to classify query shapes. Specifically, we estimate a nonparametric, square-root wavelet density on the low-order eigenvectors of the LBO, capturing the rich geometric structure of 2D and 3D shapes with very minimal pre-processing requirements. By estimating 3D, square-root wavelet densities on each shape's eigenspace (LBO-Shape densities), both 2D and 3D shapes become identifiable with the unit hypersphere. Leveraging the hypersphere's simple geometry, our avant-garde model-to-mean indexing scheme computes the intrinsic Karcher mean for each shape category, and then uses the closed-form distance between a query shape and the means to assign labels. In 2D, the need for burdensome preprocessing like extracting closed curves along with topological and equal point set cardinality requirements are eliminated. Similarly, in 3D we gain isometry invariance and rid ourselves of the need for superfluous feature extraction schemes. The extensive experimental results demonstrate that our approach is competitive with the state-of-the-art in 2D/3D shape matching algorithms.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

The widespread use of 2D and 3D shape models have spawned a staggering amount of search applications in which the models must be stored, indexed and retrieved in a manner analogous to text. To be effective on a large scale, these applications must satisfy requirements such as: compressible shape representations, rich discrimination capabilities with categorical flexibility, transformation invariance, computational scalability, etc. Our proposed shape classification framework satisfies many of these sought-after characteristics, like supporting multiresolution sparse representations, inherent robustness to isometric transformations, and having a simple closed-form similarity measure. We achieve these notable advantages as a result of estimating a nonparametric probabilistic model on the eigenspace the shape's Laplace–Beltrami operator (LBO) (dubbed *LBO-Shape densities*). Through our chosen expansion of a nonparametric density in a wavelet basis, each density be-

comes a point on a unit hypersphere, whose geometry we leverage to calculate intrinsic statistics directly on the manifold of the shape representation. The similarity between shapes is computed using the closed-form distance between the densities on the hypersphere.

Though there exists a variety of shape representation frameworks ranging from working directly with unstructured point-sets (Chui and Rangarajan, 2000; Grünwald, 2005) to weighted graphs (Siddiqi et al., 1998), curves, surfaces and other geometric models (Srivastava et al., 2007), we elect to use a density function representation. With advantages such as elimination of topology-based preprocessing (e.g. curve or surface extraction) or curtailing explicit correspondence discovery (Chen et al., 2010), the representation of geometric shapes as probabilistic distributions has yielded state-of-the-art performance in a myriad of shape analysis tasks; spanning the gamut from registration (Chui and Rangarajan, 2000; Jian and Vemuri, 2011; Peter et al., 2008; Rangarajan et al., 1997) to metric learning and shape classification (Liu et al., 2012; Moyou et al., 2014; Moyou and Peter, 2012; Thakoor et al., 2007). There are mathematical and computational benefits in working with scalar fields (density functions) that have encoded the

* Corresponding author.

E-mail addresses: mmoyou@my.fit.edu (M. Moyou), ihouk2002@my.fit.edu (K. Eddy Ihou), apeter@fit.edu (A.M. Peter).

underlying domain. Other benefits of a probabilistic representation include: elimination of equal, point-set cardinality among shapes and the inherent robustness to noise and localization error of the shape features and landmarks.

The benefits of this probabilistic representation motivated us to extend the shape modeling approach first detailed in [Peter et al., \(2008\)](#) and extended by [Moyou and Peter \(2012\)](#). The common theme between our 2D/3D shape classification methodologies is the use of the square-root wavelet density estimator ([Peter and Rangarajan, 2008](#)) and its associated hypersphere geometry. The 2D approach outlined in [Moyou and Peter \(2012\)](#) required the use of a 2D square-root wavelet density estimator, which was extended in [Moyou et al. \(2014\)](#) to three dimensions, making it the first implementation of a 3D square-root wavelet density estimator. Our 2D shape matching capability emanates from the use of the graph Laplacian (a discretization of the LBO) on the unordered 2D shape points. This present work aggrandizes the approach in [Moyou et al. \(2014\)](#) to subsume 2D shape matching—providing a single unified framework for 2D and 3D shapes—and thoroughly evaluates it on several 2D/3D shape benchmarks. A key component of our approach, as previously demonstrated in these works, is the rich modeling power afforded by the non-parametric wavelet density estimator.

The goal of wavelet density estimation is to utilize observed data samples $X = \{x_i\}_{i=1}^N$ to discover the best approximation to the underlying density—where the density is expanded in a wavelet basis. The fundamental idea behind wavelet density estimation (for 1D data) is to represent a density p as a linear combination of wavelet bases

$$p(x) = \sum_{j_0,k} \alpha_{j_0,k} \phi_{j_0,k}(x) + \sum_{j \geq j_0,k} \beta_{j,k} \psi_{j,k}(x), \quad (1)$$

where $x \in \mathbb{R}$, $\phi(x)$ and $\psi(x)$ are the *scaling* (a.k.a. father) and *wavelet* (a.k.a. mother) basis functions respectively, with $\alpha_{j_0,k}$ and $\beta_{j,k}$ being the scaling and wavelet basis function coefficients; the j -index represents the current resolution level and the k -index is the integer translation value of the basis function. The uniqueness inherent in our methodology stems from the fact that we seek to expand the $\sqrt{p(x)}$ in a wavelet basis, and recover the original density as $(\sqrt{p(x)})^2$ —more specifically, we seek the coefficients of $(\sqrt{p(x)})^2$ expanded in a wavelet basis. Each density is represented by its vector of coefficients which, as detailed in [Peter and Rangarajan \(2008\)](#), serve as coordinates of a point on a unit hypersphere. In the present context, the data samples come from the eigenvector embedding coordinates of the discretized LBO.

After estimating the coefficients of the wavelet densities, we compute a mean density by calculating the Karcher mean ([Karcher, 1977](#)) for each shape category, directly on the hypersphere. The collection of mean densities, serves as the new index (mean index) for classifying shapes. We then compute the distance from the query shape to each entry in the mean index, and the query shape is assigned the category label of the closest mean density. A visual representation of the LBO-Shape density framework is shown in [Fig. 1](#).

This divergence from the customary model-to-model comparison ([Bronstein et al., 2011](#); [Khouri et al., 2012](#); [Li and Hamza, 2013](#)), creates a much more efficient index by only having to search within the labels of the mean densities to classify query shapes. The index is further fortified because all the information of the shape category is contained within a single model, hence removing the majority of the misclassifications due to minute differences in the shapes. As demonstrated in our experimental results, our enhanced classification process achieves notable scores on several bench-marked datasets.

2. Related work

The work ([Mingqiang et al., 2008](#)) effectively covers the myriad of 2D shape matching techniques that exist in the literature, where the almost de facto shape representation model is the use of closed curves ([Felzenszwalb and Schwartz, 2007](#); [McNeill and Vijayakumar, 2006](#)). The extraction of these curves are extremely sensitive to noise and occlusions—their adoption often imposes undesirable topological restrictions and point set cardinality requirements on the shapes. The authors of [Cao et al. \(2011\)](#) propose a MCMC (Markov Chain Monte Carlo) based algorithm where shape contours are represented by a set of ordered landmark points. A posterior distribution is used to select a sequence of landmark points that are compared using the Procrustes distance. Their prior distribution awards the matching of longer segments more than smaller ones, hence increasing matching accuracy. [Srivastava et al. \(2005\)](#) proposed a differential-geometric treatment of planar shapes using closed curves which includes hierarchical clustering, learning of probabilistic models from clusters of shapes and varying the probabilistic model to retrieve shapes. Our method is more effective than [Cao et al. \(2011\)](#); [Felzenszwalb and Schwartz \(2007\)](#); [McNeill and Vijayakumar \(2006\)](#); [Srivastava et al. \(2005\)](#) because we use all of the unlabeled points to construct a graph instead of noise-sensitive extracted curves that only use a subset of the points. This allows us to encode the shape's detailed internal structure, providing more accurate representations of shapes.

[Bimbo and Pala \(2006\)](#) consider the existing shapes as a group and study their similarity measures to the query shape in a graph structure. For some given similarity measure, a new similarity is learned iteratively through graph transduction. Graphs are built on subsets of shapes, where all the shapes from the query class are assumed to be in the set, and the geodesic distance on the graph determines similarity. The graph construction in our approach is used to develop a spherical manifold structure in which the similarity measure is in closed-form and computationally efficient. More importantly, we guarantee the efficient comparison of the query shape to *all* models through the formation of mean shapes per category, allowing us to achieve higher accuracies.

The 3D shape classification problem has been extensively investigated in the literature and is well documented by [Tangelder and Veltkamp \(2008\)](#). Early on, one of the most popular spectral-based techniques was Reuter et. al.'s Shape-DNA ([Reuter et al., 2006](#)) that used the truncated spectrum (sequence of eigenvalues) of the LBO as an isometry invariant shape descriptor. This motivated [Rustamov \(2007\)](#) to create the Global Point Signature (GPS) embedding, where each eigenvector of the LBO is scaled with the square root of its corresponding eigenvalue (analogous to the scaling employed for eigenvectors of operators on finite dimensional vectors spaces). Whereas Rustamov uses up to 25 eigenvectors to construct histograms of pairwise distances between points, we use only three low order eigenvectors and directly estimate a robust, non-parametric density function on the eigenspace, and our approach provides a unified framework for 2D and 3D matching. It is important to note that the use of density functions (approximated by histograms) for content-based retrieval was motivated by the seminal work in [Rubner et al. \(2000\)](#) through the use of the Earth's Mover's Distance (EMD). The EMD is also known as the Wasserstein metric (first order) or Mallow's distance ([Levina and Bickel, 2001](#)). The use of the EMD to compare 1D histograms is very efficient but requires linear programming optimization algorithms to solve for higher dimensions. This is different to our spherical closed-form distance measure which efficiently compares density functions.

[Sun et al. \(2009\)](#) developed the Heat Kernel Signature (HKS) by observing the heat diffusion over time along the surface. Build-

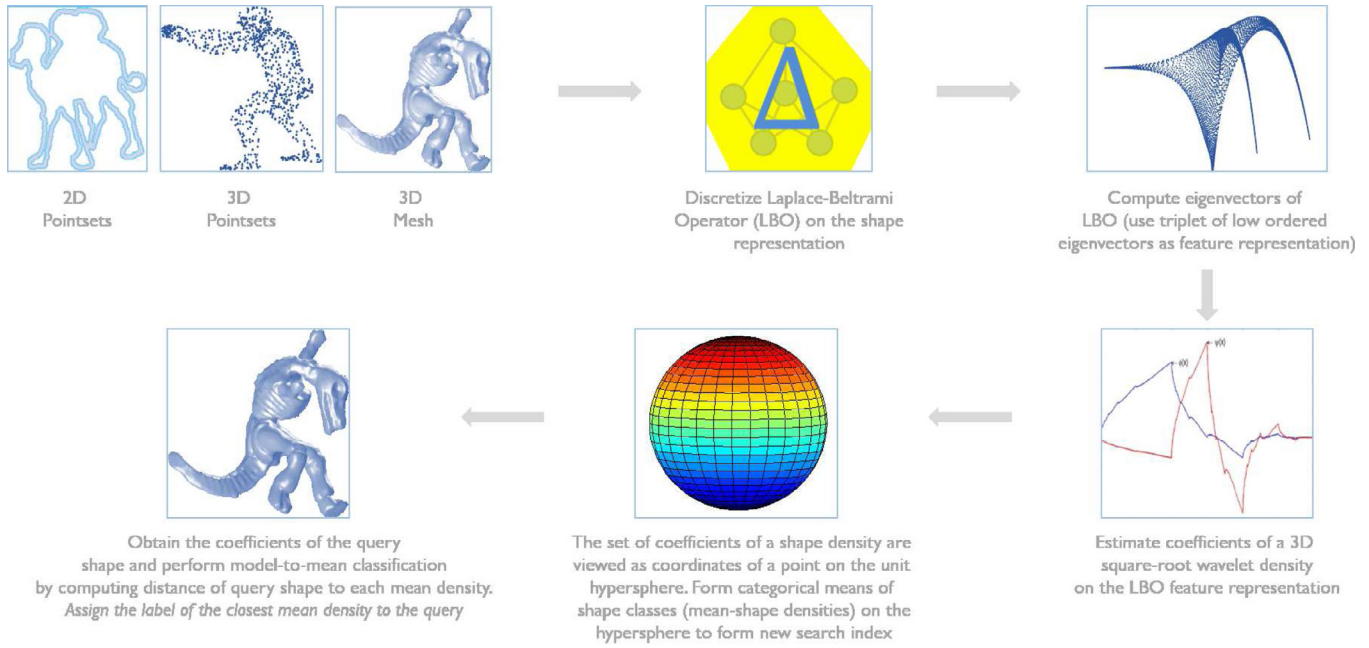


Fig. 1. LBO-Shape density overview.

ing upon the HKS (Sun et al., 2009), Bronstein et al. developed the scale invariant heat kernel signature (SI-HKS) (Bronstein and Kokkinos, 2010) and Shape Google (Bronstein et al., 2011). In the Shape Google framework, they algorithmized the practical use of the HKS for retrieval applications by incorporating the common bag-of-words framework from text retrieval.

The approach most closely related to ours is Li and Hamza (2013), who proposed a multiresolution spectral signature using a graph wavelet framework. Their shape signature is the set of coefficients at a particular resolution level generated from the cubic spline wavelet generating kernel. Shapes are compared using intrinsic spatial pyramid matching that involves the use of the level sets of the second eigenvector of the LBO as cuts to perform surface partition. By using an eigenvector *triplet* of a shape's LBO, our wavelet coefficient signature inherently contains more comprehensive geometric information compared to using only the second eigenvector in Li and Hamza (2013). In addition, the LBO-Shape density technique that we employ results in a manifold representation that promotes aggregation of shape category information into a single model, resulting in smaller indexes for classification.

The rest of this paper is outlined as follows. Section 3 describes the theoretical and practical details of our 2D/3D shape matching framework. In Section 4, we provide extensive experimental results to validate the approach's favorable performance on a variety of evaluation metrics and datasets. Finally, Section 5 provides a summary and future research.

3. Theory

Our classification methodology can be partitioned into three major components. The first component requires discretization of the LBO on the shape manifold and then performs an eigendecomposition of this discretized LBO to produce our eigenshapes. The second, 3D square-root wavelet density estimation, moves from the point-set topology of the eigenshape to a parametric model based on the coefficients of the density estimator, which also maps each LBO-Shape density to a point on the unit hypersphere. The third component, leverages the manifold structure of the shape densities to produce class-specific prototypes for the differing shape categories, enabling efficient classification of either 2D or 3D shapes.

To begin with, the original 2D shapes are simply unlabeled points sets, without requirements on the cardinality of the points (see Fig. 2). The input 3D shapes are both triangular mesh models and point sets. Since we need discrete approximations to the LBO, we utilize the graph Laplacian when working with 2D and 3D point sets, and the cotangent approximation for 3D triangular meshes. It is worth noting that discretizing the graph Laplacian on a shape and using its associated spectral embedding was originally developed by Belkin and Niyogi (2003), better known as Laplacian Eigenmaps. In our method, we refer to this embedding space as the "LBO-Shape space" because we utilize two different discretizations of the LBO in the same framework. The geometric information gained from the eigendecomposition of these LBO discretizations can be utilized through density estimation.

Our density estimation model is a fully nonparametric, data-driven estimator emanating from the expansion of the square-root of the 3D density function in a wavelet basis. Since we adopt a nonparametric philosophy, we maximize the information harnessing capability of our method, and eliminate typical distributional assumptions present in other density estimation techniques. Even more importantly, the choice of basis inherently provides a multiresolution representation enabling sparse storage for large-scale classification applications.

This wavelet density estimation procedure provides us with a hypersphere geometry where the shape similarity is measured via the simple closed-form arc length distance. The geometry can also be leveraged to learn a single representative model per shape category—resulting in a more efficient indexing structure for classification.

Classification of 2D shapes based on the LBO-Shape density technique begins with constructing a graph on the unordered shape points, followed by an eigenvector computation of its graph Laplacian. For each shape, the coefficients of a 3D square-root wavelet density (LBO-Shape density) are estimated given a triplet of the shape's low order eigenvectors (the eigenshapes formed from the triplets of eigenvectors are shown in Fig. 2). For each category, the Karcher Mean (Karcher, 1977) of shape densities is computed on the hypersphere—creating a more compact index (called the mean index), where each entry is the prototype density of a

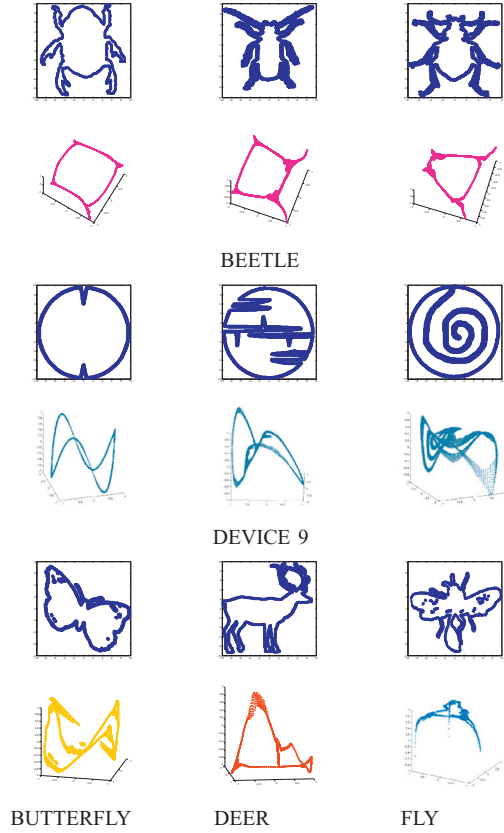


Fig. 2. Eigenshapes of various 2D models from the MPEG 7 dataset formed from the eigenvectors the graph Laplacian. The original 2D shapes are shown in rows 1,3 and 5 in blue. The image below each original shape is its corresponding eigenshape, which are formed using the eigenvector triplet (1,2,5). To match 2D shapes we estimate wavelet densities directly on these eigenshapes (*LBO-Shape densities*). Notice, that shapes within a category have similar eigenshapes whereas across categories the eigenshapes are distinctly different. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

category, and finally shapes are classified based on the minimum spherical distance to the mean index.

Our LBO-Shape density approach follows an analogous path for 3D shapes, we first compute the low order eigenvectors of the shape's LBO (see Fig. 3)—the only aspect changed from 2D is that we have the option to use either the graph Laplacian or the cotangent approximation of the LBO. Again, a triplet of eigenvectors is used to estimate a 3D square-root wavelet density for each shape. After which, the mean index is generated by executing the Karcher mean algorithm (Karcher, 1977) on the estimated densities per shape category. The minimum spherical distance of a shape density to the mean index serves as a classification criterion for the query shapes.

We now expand in detail the major components of the LBO-Shape density retrieval framework, which involves: the theoretical underpinnings of the graph Laplacian in 2D and 3D, the cotangent approximation of the LBO in 3D and all of their respective low order eigenvectors; estimating 3D square-root wavelet densities on the eigenspaces of shapes; computing the Karcher mean on the hypersphere of wavelet densities; and finally, the classification technique based the spherical distance to the mean index.

3.1. Laplace–Beltrami operator

The Laplace–Beltrami operator generalizes the Laplacian of Euclidean spaces to Riemannian manifolds. Consider a function f to

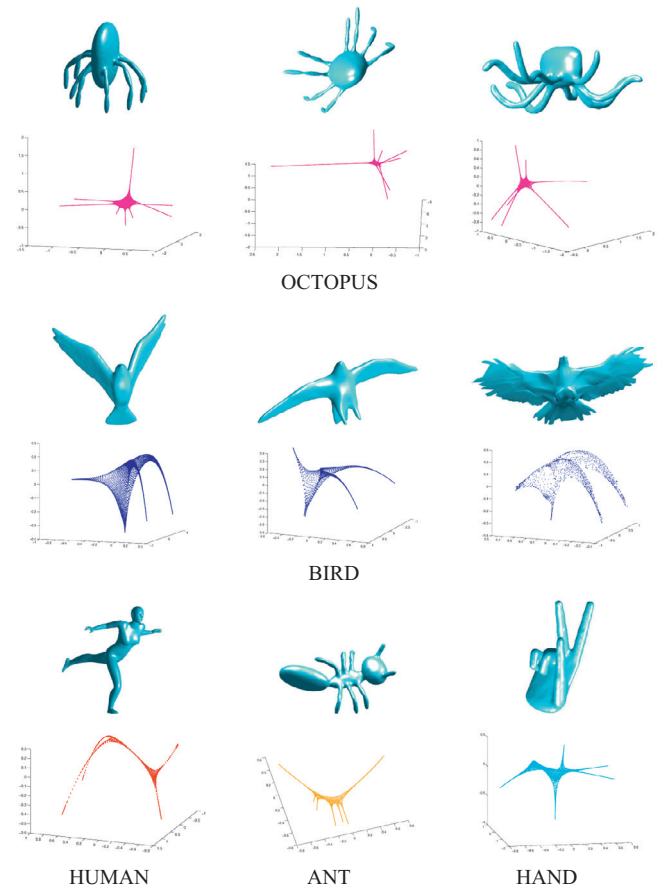


Fig. 3. Eigenshapes of various 3D shape models (in light blue) obtained from the cotangent approximation to the LBO. The eigenshapes are formed using the eigenvector triplet (1,2,5). These are matched using the same technique as in 2D. Notice, that shapes within a category have similar eigenshapes whereas across categories the eigenshapes are distinctly different. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be a C^2 real-valued function defined on a smooth Riemannian manifold M with metric tensor g . The coordinate-free Laplace–Beltrami operator Δ is defined as

$$\Delta f = -\operatorname{div}(\operatorname{grad} f), \quad (2)$$

where div and grad are the divergence and gradient on the manifold M (Isaac, 1984). In local coordinates, this reduces to

$$\Delta f = \underbrace{\frac{1}{\sqrt{|g|}} \sum_i \partial_i \sqrt{|g|}}_{\text{divergence}} \underbrace{\sum_j g^{ij} \partial_j f}_{\text{gradient}}. \quad (3)$$

where g^{ij} denotes the elements of the inverse of g .

Solving the standard *eigenvalue* problem

$$\Delta v = -\lambda v, \quad (4)$$

for λ the eigenvalue of Δ and v the eigenvector corresponding to λ , provides the necessary tools for a variety of geometric analysis on the manifold. (Note: For the LBO, $\lambda = 0$ is always an eigenvalue for which its corresponding eigenvector is constant and hence discarded in most applications, including ours). The Laplace–Beltrami operator over a compact manifold S is bounded and symmetric positive semidefinite; its set of eigenvalues are non-negative real numbers and its set of eigenvectors are countable (Zeng et al., 2012). In Jones et al. (2008), it was proven that the LBO eigenvectors are theoretically a good local parametrization for Riemannian

nian manifolds, affirming their use as coordinates in our subsequent density estimation step. For shape matching, the eigenvectors have several other useful properties (Rosenberg, 1997):

1. The eigenvectors are intrinsic to the surface and invariant under isometric deformations.
2. The set of eigenvectors forms a complete orthonormal basis for the \mathbb{L}^2 space of functions on the manifold.
3. Two manifolds are isometrically deformable into each other if and only if their Laplacian operators have the same eigenvalues and eigenvectors.

For shape analysis, the eigenvectors of the LBO capture the interaction between the topology and geometry at a fundamental level (Levy, 2006). Intuitively, the action of the LBO can be seen as spreading a function over small neighborhoods on shapes with arbitrary topology; the eigenvectors give insight into the interaction of these local neighborhoods and their role in the emergence of global properties.

For computational applications, one has to discretize the LBO which results in a finite dimensional operator. Several discretization schemes exist, e.g. graph Laplacian (Fan, 1997) and point-based Laplacian (Liu et al., 2012). The authors in Xu (2004) proposed other LBO discretization schemes and established the theoretical basis for their convergences. Wardetzky et al. (2007) proved that the diversity of existing discrete Laplace operators resulted from their inability to satisfy all the natural properties of their continuous counterparts. The authors rigorously prove the non-existence of a “perfect” discrete Laplacian satisfying all of the properties such as: symmetry, local support, linear precision, positive definiteness, etc. Even though a perfect discrete Laplacian does not exist, as mentioned before, the graph Laplacian is the most suitable LBO approximation for our 2D shape classification framework given that we assume a point set representation.

3.2. Graph Laplacian

We begin this section with some relevant definitions and standard background about graphs. A graph G is a pair (V, E) , where V is a finite set of vertices and E is a set of connections (edges) between the vertices. The size of the graph is defined as the number of vertices. For two vertices $u, v \in V$, the connection between them is called an edge $e = (u, v)$. An edge weighted graph $G = (V, E)$ has weights $w(e) \in \mathbb{R}$ associated with each edge $e \in E$. Graphs whose edges have a weight of one are known as unweighted graphs. A graph is simple if it does not contain any self-loops or multiple edges between vertices, and thus its edge set only contains distinct pairs. The graphs used in the LBO-Shape density paradigm are all simple.

The *adjacency matrix* A of a graph $G = (V, E)$ is a $|V| \times |V|$ matrix whose (u, v) element is given as:

$$A(u, v) = \begin{cases} 1 & \text{if } (u, v) \in E \\ 0 & \text{otherwise} \end{cases}$$

D is the diagonal matrix of vertex degrees whose elements are $D(u, u) = \sum_{v \in V} A(u, v)$. One popular method of constructing graphs is through an ϵ -neighborhood. For the adjacency matrix, two vertices are connected if the distance between them is less than some ϵ distance. The lengths of the edges can be used as weights to construct a weighted adjacency matrix.

3.2.1. Graph Laplacian construction

To construct graphs in 2D and 3D we assume that the shape points have been extracted in 2D and are available from the mesh representations in 3D. The first step is to fit the shapes into a square grid (a cube in 3D), which ensures that the ϵ value chosen to threshold distances can apply to all shapes. Without this

renormalization, a single ϵ value may not apply to all the shapes and different laplacian embeddings may result (see Belkin and Niyogi, 2003 for more details). The second step is to compute all the pairwise distances between the points in the shape and threshold the distances using ϵ to form the adjacency matrix A , which can be used to form the Laplacian.

The *Laplacian matrix* of G or *graph Laplacian* is a symmetric positive semidefinite matrix given as $L = D - A$. The spectral decomposition of the graph Laplacian is given as:

$$L\phi = \lambda\phi,$$

where λ is an eigenvalue of L with a corresponding eigenvector ϕ . The eigenvalues of the graph Laplacian are non-negative and constitute a discrete set. The spectral properties of L are used to embed the points into a lower dimensional space, and gain insight into the geometry of the shape (Isaacs and Roberts, 2011). The LBO-Shape density technique possesses no equal point set cardinality requirements and topological constraints because we use the unordered shape points to construct undirected graphs. The number of points in each shape varies, therefore the eigenvectors of their corresponding graph Laplacians will be of different lengths. By estimating wavelet densities on these eigenvectors we eliminate the point-set cardinality requirements (see Section 3.5). Next, we look at the cotangent approximation of the LBO for 3D shapes.

3.3. Cotangent approximation

The cotangent Laplacian approximation uses a linear finite element method (FEM) to discretize the LBO and was first detailed in Dziuk (1988); Pinkall et al. (1993). To compute the solution to (4), we verify it using its weak form,

$$\langle \Delta v, \varphi_i \rangle = -\lambda \langle v, \varphi_i \rangle \quad \forall i \quad (5)$$

for some test function φ_i under the \mathbb{L}^2 inner product. Finding the eigenvectors of the LBO with linear FEM amounts to solving the *generalized eigenvalue problem*

$$A_{\text{cot}} v_i = -\lambda B v_i, \quad (6)$$

with

$$A_{\text{cot}}(i, j) := \begin{cases} \frac{\cot \kappa_{i,j} + \cot \xi_{i,j}}{2} & \text{edge } (i, j) \\ -\sum_{k \in N(i)} A_{\text{cot}}(i, k) & i = j \end{cases}$$

$$B(i, j) := \begin{cases} \frac{|t_1| + |t_2|}{12} & \text{edge } (i, j) \\ \frac{\sum_{k \in N(i)} |t_k|}{6} & i = j, \end{cases} \quad (7)$$

where $|t_i|$ is the area of triangle t_i . The variables t_1 and t_2 are the triangles that share edge (i, j) ; B is the mass matrix; A_{cot} is the stiffness matrix with cotangent weights, $\kappa_{i,j}$ and $\xi_{i,j}$ are the angles opposite to the edge (i, j) (Reuter et al., 2009). Through this formulation, the weighted inner product induced by B , i.e. $\langle f, g \rangle_B = f^T B g$ for $f, g \in \mathbb{L}^2$, generalizes the \mathbb{L}^2 inner product (i.e. $B := I$); it is intrinsic to the surface on which it is defined; and is adapted to the local sampling of M through the variation of triangle areas (Patané, 2013). The mass matrix B also accurately encodes the geometry of the input surface through the area of its triangles and leads to FEM distances having higher robustness to topological and scale changes, irregular sampling and noise. Note that this formulation of the LBO only works on closed manifolds which was the focus of this initial research effort; extensions to other LBO discretizations will be pursued in future research. It is well known that the eigenvectors of the decomposition of a matrix are recovered up to a sign factor of ± 1 , and any of these methods (Caelli and Kosinov, 2004; Sang et al., 2000; Umeyama, 1988; White and Wilson, 2007) can be used to sufficiently address this problem.

Eigenvectors corresponding to smaller eigenvalues correlate to lower frequency mode properties of the shape, and larger

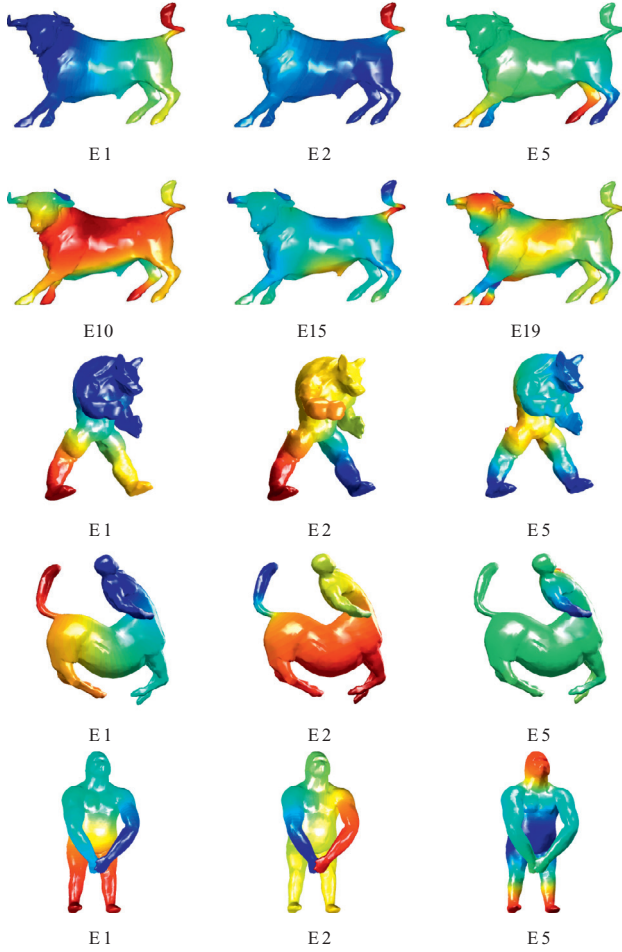


Fig. 4. Example eigenvectors of a bull (Shape COSEG, Four Leg category), armadillo, centaur, and gorilla (SHREC'11). The subtitles indicate which eigenvector is plotted on the shape, for instance E10 is the 10th eigenvector. The low order eigenvectors represent the lower frequency modes of the shape, and are more robust to noise and deformations.

eigenvalue-eigenvector pairs capture high frequency modes of the surface—see Fig. 4 for a visualization of these characteristics. In our extensive empirical evaluations, we have found that using triplet combinations from the 10 lowest order eigenvectors are sufficient to discriminate among the various shape categories. Once the desired triplet of eigenvectors are selected, our objective now becomes to estimate the trivariate distribution $p(v_1, v_2, v_3)$ on the eigenvectors in a wavelet basis. Before wavelet density estimation is discussed, we first give some background about wavelets.

3.4. Wavelets and multiresolution analysis

Wavelet bases have the desirable property of being able to approximate functions from \mathbb{L}^2 , the space of square-integrable functions. For any function $f \in \mathbb{L}^2$ and a starting resolution level j_0 , f 's representation in a wavelet basis is given as

$$f(x) = \sum_{j_0, k} \alpha_{j_0, k} \phi_{j_0, k}(x) + \sum_{j \geq j_0, k} \beta_{j, k} \psi_{j, k}(x), \quad (8)$$

where

$$\begin{aligned} \phi_{j_0, k}(x) &= 2^{j_0/2} \phi(2^{j_0}x - k) \\ \psi_{j, k}(x) &= 2^{j/2} \psi(2^jx - k) \end{aligned} \quad (9)$$

are scaled and translated versions of the father $\phi(x)$ and mother $\psi(x)$ wavelets. The crucial idea in multiresolution theory is decom-

posing the \mathbb{L}^2 function space into a sequence of nested subspaces V_j , $j \in \mathbb{Z}$ such that

$$\dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \dots, \quad (10)$$

where the subspaces are non-intersecting i.e. $\bigcap V_j = \{0\}$ and $\bigcup V_j = \mathbb{L}^2$. The resolution increases as $j \rightarrow \infty$ and decreases as $j \rightarrow -\infty$. For any particular resolution level j , at the $j+1$ th level, we have the following relationship:

$$V_j \oplus W_j = V_{j+1}, \quad (11)$$

where W_j is a space orthogonal to V_j , i.e. $V_j \cap W_j = \{0\}$. The spaces V_0 and W_0 are intimately related to the father wavelet $\phi(x)$ and the mother wavelet $\psi(x)$, respectively. The father wavelet $\phi(x)$ and its integer translates form a basis for V_0 . The mother wavelet $\psi(x)$ and its integer translates span W_0 . The spaces decompose the function into its smooth and detailed components analogous to viewing a function at different scales, where at each scale, the function possesses both a low pass and a high pass version.

We assume that $\phi(x)$ and $\psi(x)$ in conjunction with their scaled and translated versions form orthogonal bases for their respective spaces. Based on these assumptions, the standard technique for calculating coefficients for (8) is to use the inner product of the space, the \mathbb{L}^2 inner product in this case. For example, the coefficient $\alpha_{j_0, k}$ is obtained by

$$\alpha_{j_0, k} = \langle f, \phi_{j_0, k} \rangle = \int f(x) \phi_{j_0, k}(x) dx. \quad (12)$$

This projection paradigm for coefficient estimation is employed for most of the existing wavelet-density estimation techniques. Since we are estimating a *density function*, the \mathbb{L}^2 inner product yields an interesting result. The coefficients for (1) are calculated as the expectation of the random variable evaluated at the corresponding basis function:

$$\alpha_{j_0, k} = \int p(x) \phi_{j_0, k}(x) dx = \mathbb{E}[\phi_{j_0, k}(x)], \quad (13)$$

where \mathbb{E} is the expectation operator. Given N samples, the scaling function coefficient estimation of $\alpha_{j_0, k}$ is approximated by the sample average

$$\alpha_{j_0, k} = \frac{1}{N} \sum_{i=1}^N \phi_{j_0, k}(x_i). \quad (14)$$

Similarly the wavelet function coefficients $\beta_{j, k}$ are estimated as

$$\beta_{j, k} = \frac{1}{N} \sum_{i=1}^N \psi_{j, k}(x_i).$$

Our density estimation technique requires evaluating $\phi(x)$ and $\psi(x)$ at various domain points in their support region; however, wavelets do not exist in closed-form and are generated numerically, the reader is referred to Peter and Rangarajan (2008) for more details. With the knowledge of wavelets and density estimation now covered, we discuss how to construct a density function on the shape's eigenspace using a 3D square-root wavelet formulation.

3.5. Square-root wavelet density estimation

The excellent convergence properties of wavelets (Donoho et al., 1996) coupled with their inherent ability to model a large class of functions—especially those with sharp transitions—make them a prime candidate for modeling density functions. The utility of a wavelet density estimation approach for shape recognition in 2D has been previously shown by Peter et al. (2008). One of the main challenges faced by the majority of basis function-based

non-parametric estimators is of generating a bonafide density estimate, i.e. one that satisfies non-negativity and unit integrability $\int p(x)dx = 1$ constraints. One solution is to first estimate $\sqrt{p(x)}$ and then obtain the desired density as $(\sqrt{p(x)})^2$.

In our LBO-shape density approach we must estimate a density function of the eigenvectors of the LBO. An interesting point to note is that the histograms of eigenfunctions are invariant to node permutation, i.e., invariant to the order in which the components of the eigenvectors are considered (Isaacs and Roberts, 2011). This is also true for the case of the wavelet density estimator since it belongs to the same class of non-parametric estimators. Therefore the wavelet density signature can be viewed as an invariant signature of an eigenvector. We use this eigenvector invariance to compute a 3D square-root density function represented in a wavelet basis.

The 3D wavelet expansion of the square-root of the density is given by

$$\sqrt{p(\mathbf{x})} = \sum_{j_0, \mathbf{k}} \alpha_{j_0, \mathbf{k}} \phi_{j_0, \mathbf{k}}(\mathbf{x}) + \sum_{j \geq j_0} \sum_{\mathbf{k}} \beta_{j, \mathbf{k}}^w \psi_{j, \mathbf{k}}^w(\mathbf{x}), \quad (15)$$

where $\mathbf{x} \in \mathbb{R}^3$, J is the stopping scale level for the multiresolution decomposition and $\mathbf{k} = (k_1, k_2, k_3) \in \mathbb{Z}^3$ is a multi-index that represents the spatial translation location of the basis in three dimensions. The translation range of \mathbf{k} can also be computed from the span of the data and the basis functions support size. Recall, that we are estimating the density function on a triplet of eigenvectors \mathbf{v} of the LBO, so $\mathbf{x} = (v_1, v_2, v_3)$ serve as the random variables. To obtain three-dimensional basis functions, we use tensor product constructions of the father's and mother's one dimensional bases which results in eight combinations, i.e.

$$\begin{aligned} \phi_{j_0, \mathbf{k}}(\mathbf{x}) &= 2^{\frac{3j_0}{2}} \phi(2^{j_0}x_1 - k_1) \phi(2^{j_0}x_2 - k_2) \phi(2^{j_0}x_3 - k_3) \\ \psi_{j, \mathbf{k}}^1(\mathbf{x}) &= 2^{\frac{3j}{2}} \phi(2^jx_1 - k_1) \psi(2^jx_2 - k_2) \psi(2^jx_3 - k_3) \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \psi_{j, \mathbf{k}}^6(\mathbf{x}) &= 2^{\frac{3j}{2}} \psi(2^jx_1 - k_1) \psi(2^jx_2 - k_2) \phi(2^jx_3 - k_3) \\ \psi_{j, \mathbf{k}}^7(\mathbf{x}) &= 2^{\frac{3j}{2}} \psi(2^jx_1 - k_1) \psi(2^jx_2 - k_2) \psi(2^jx_3 - k_3). \end{aligned} \quad (16)$$

Given a set of 3D samples (eigenvector triplet), the objective becomes to estimate the basis coefficients $\rho = \{\alpha_{j_0, \mathbf{k}}, \beta_{j, \mathbf{k}}^w\}$, and then use them to reconstruct the density expansion (15). The full multiresolution model in Eq. (15) supports thresholding of the $\beta_{j, \mathbf{k}}^w$ detail coefficients which can produce very accurate sparse representations of the density functions. For our initial investigations presented in this paper, we have elected not to explore this aspect; hence, we only use scaling function expansion of the densities. This still yields the same density estimate provided appropriate selection of the starting j_0 resolution level (Peter and Rangarajan, 2008), which is our primary interest for determining shape similarity. The work in Peter and Rangarajan (2008) provides an efficient maximum likelihood method to estimate these coefficients, with fast convergence. However, it only developed 1D and 2D estimators, here we have extended the optimization framework for 3D density estimation. For 3D estimation, the optimization framework outlined in Peter and Rangarajan (2008) requires very minimal changes to address the increase in tensor product basis functions, which lead to potentially 8 coefficients at each location in the 3D lattice. (Since this bona fide 3D wavelet density estimator can be independently applied in many other applications, we plan to make the source code available in the near future¹). Using the

details of our density estimator just covered, we explore the resulting manifold structure and how this structure promotes a more efficient classification scheme.

3.6. The hypersphere geometry of wavelet densities

When using an orthogonal basis expansion—like the Daubechies, Coiflet, or Symlet wavelets—the square-root estimation model implicitly provides a geometric characterization of the space of densities. This stems from the unit integrability requirement of probability density functions, which translates to a constraint on the basis function coefficients, i.e. $\int (\sqrt{p(\mathbf{x})})^2 d\mathbf{x} = 1$ implies

$$\sum_{j_0, \mathbf{k}} \alpha_{j_0, \mathbf{k}}^2 + \sum_{j \geq j_0} \sum_{\mathbf{k}} \sum_{w=1}^8 (\beta_{j, \mathbf{k}}^w)^2 = 1. \quad (17)$$

This leads to the interpretation of each density function being a point on the unit hypersphere S^{d-1} . More specifically, the set of coefficients ρ serve as coordinates of a point on this $d-1$ hypersphere, where d is the cardinality of the coefficient set $d = |\rho|$.

The hypersphere geometry of the densities can be more rigorously justified when we analyze the $\sqrt{p(x)}$ representation under the theoretical basis of information geometry (Amari and Nagaoka, 2001; Peter and Rangarajan, 2009). In this context, the Fisher information matrix (FIM) serves as the metric tensor on the manifold of a parametric family of distributions. One of the algebraic forms of the FIM is given by

$$g_{u,v} = 4 \int \frac{\partial \sqrt{p(x|\Theta)}}{\partial \Theta^u} \frac{\partial \sqrt{p(x|\Theta)}}{\partial \Theta^v} dx, \quad (18)$$

where $\Theta = \{\theta^1, \dots, \theta^m\}$ denotes the parameters of the distribution and u and v indicate the row and column index, i.e. for a family with m parameters the FIM is $m \times m$. Under an orthonormal expansion of $\sqrt{p(x|\Theta)}$, Eq. (18) reduces to the canonical metric tensor of a unit hypersphere embedded in an $m+1$ dimensional Euclidean space. Rather than use the metric tensor to intrinsically compute geodesics on the hypersphere (an undertaking which would require us to parametrize the manifold), we can accomplish the same computation by realizing that the constraint $\sum_{i=1}^{m+1} (\theta^i)^2 = 1$ also implies the unit hypersphere geometry. Hence, closed-form geodesics distances can be simply computed using the usual angle measure between two unit vectors. Such is the case in our framework where $\sqrt{p(x|\Theta)}$ has been expanded in an orthonormal wavelet basis with the coefficients of the expansion serving as the parameters of the density, i.e. $\Theta = \{\alpha_{j_0, \mathbf{k}}, \beta_{j, \mathbf{k}}^w\}$, previously denoted as ρ .

3.7. Intrinsic means on the wavelet hypersphere

Thus far, we have discussed our approach to representing shapes as density functions synonymous with points on a unit hypersphere through the process of wavelet density estimation. The eigenspace spanned by three low order eigenvectors of the shapes LBO approximation serve as the feature space in which these wavelet densities were computed. Given exemplars from a particular shape category, we can obtain a prototype representation of that category by computing a mean model. In the proposed framework, this notion translates to computing a mean density function from the LBO-Shape densities of a particular category. Keep in mind, with our framework, both 2D and 3D shape matching require estimation of 3D densities, resulting in similar hypersphere representations of the densities. Hence, they share the same classification mechanics—a unique and unifying property of LBO-Shape densities.

¹ ICE Lab Software: <http://research2.fit.edu/ice/?q=node/26>

Since the densities are points on the manifold (hypersphere), obtaining a mean density function requires us to compute the generalized Karcher mean (Karcher, 1977). To execute this intrinsically on the manifold, we employ the Exponential (Exp) and Logarithm (Log) maps on the manifold (available as analytic formulas for the hypersphere), and implement the simple optimization procedure detailed in Algorithm 1 (Pennec, 2006).

The Exp map takes a vector γ on the tangent space at $\rho_1, \gamma \in T_{\rho_1}(S^{n-1})$, and returns a point ρ_2 on the hypersphere

$$\rho_2 = \text{Exp}_{\rho_1}(\gamma) = \cos(|\gamma|)\rho_1 + \sin(|\gamma|)\frac{\gamma}{|\gamma|}. \quad (19)$$

Conversely, the Log map takes a point ρ_2 on the hypersphere and returns a vector on the tangent space at ρ_1 , by letting

$$\begin{aligned} \tilde{\rho} &= \rho_2 - \langle \rho_2, \rho_1 \rangle \rho_1 \\ \gamma &= \text{Log}_{\rho_1}(\rho_2) = \tilde{\rho} \frac{\cos^{-1}(\langle \rho_1, \rho_2 \rangle)}{\sqrt{\langle \tilde{\rho}, \tilde{\rho} \rangle}}. \end{aligned} \quad (20)$$

3.8. Shape classification using intrinsic means

Traditional shape matching algorithms are founded on the model-to-model philosophy, with lists ordered according to the decreasing similarity of the database elements with the query element. The discrimination of elements between categories hinges upon the individual shape representation, with the size of the index being proportional to the number of shapes in the database. For prodigious databases, the speed of classification is limited by having to search over large index structures. In the LBO-Shape density paradigm, we circumvent this expensive search through the use of an avant-garde model-to-mean procedure. We now explain the classification formulation using the mean index and its differences with the ubiquitous model-to-model scheme.

A paramount objective in shape classification is to devise a representation that robustly distinguishes shapes between categories, grouping the models that are most similar with each other. In our technique, we fuse the shape information within a category to produce a single prototype model; subsequently, reducing the size of the index to be proportional to the number of categories. In most databases, the number of elements is expected to dwarf the number of categories. Under this assumption, our categorically inspired mean index allows for a much more efficient labeling. Since we create prototype models of shape categories we decrease the variability of discrimination error due to the minute differences between shapes.

For the shape classification problem with multiple categories, our approach is to compute a mean density for each of the classes (mean index) using the associated densities in each class. Given a query shape we estimate its wavelet density from the corresponding eigenspace and compute the distance of the query density to each entry in the mean index using the closed-form distance on the hypersphere

$$d(\rho, \mu_i) = \cos^{-1}(\rho^T \mu_i), \quad (21)$$

where ρ is a vectorized set of wavelet coefficients for the query shape density and μ_i is the set of coefficients associated with the mean shape density of the i th class. The category label of the closest mean density is assigned to the query shape. It is worth noting, that all of our analysis is taking place intrinsically on the manifold of our shape representation, an added advantage over other methodologies that decouple the representation and matching. Next, we detail our experimental evaluations and show the utility of this approach on multiple databases.

4. Experimental results

In the LBO-Shape density approach we compute the eigenvectors of two discrete discretizations of the Laplace–Beltrami operator (graph Laplacian and cotangent approximation). Armed with the knowledge that low order eigenvectors contained useful geometric information about the shapes, this motivated us to investigate the utility of combining that information into a classification framework. Previous successful work on 3D shape matching (Moyou et al., 2014) proved that this method was indeed noteworthy, and prompted our present extension into 2D shape classification.

The specific combination of the low order eigenvectors was chosen experimentally based on the knowledge that low frequencies of the Laplace–Beltrami spectrum are related to information about the global structure and features of the input surface, while high frequencies reflect fine changes in the shapes (Ruggeri et al., 2010). To arrive at the (1,2,5) eigenvector triplet (ET), extensive 2D/3D shape matching experiments were run with numerous combinations of triplets using the first 15 eigenvectors. This was done only after conducting investigative experiments on ETs containing different combinations of low and high order eigenvectors. Note that experiments were conducted with tuples of eigenvectors, but they did not demonstrate state-of-the-art performance—this further supported the use of eigenvector triplets. Results with the (1,2,3) combination are also provided to show that different triplets can yield fairly similar results. The matching results based on only high order eigenvectors were less significant than those of the low order eigenvectors. This was due to the detailed nature of these higher order eigenvectors and their inability to model common information amongst similar shapes. The combination of low and high order eigenvectors also yielded less significant results, because once again, the higher order eigenvectors are more sensitive to the subtle changes in the shape's surface geometries, therefore shapes within a category were almost as dissimilar as shapes outside of the category.

To assess the performance of the shape matching framework we used several popular shape datasets: the MPEG-7 database (Latecki et al., 2000) and Brown dataset (Sebastian et al., 2001) for 2D, and SHREC'15 (Lian et al., 2015), SHREC'12 (Godil; and et al., 2012), SHREC'11 (Lian et al., 2011), SHREC'10 (Lian et al., 2010) for 3D (see Fig. 5). In this section, we separately outline the experimental procedure and results for the LBO-Shape density approaches in 2D and 3D.

4.1. 2D shape classification results

As mentioned above, we evaluated the 2D shape matching framework on the popular MPEG-7 CE-Shape-1 part B database (Latecki et al., 2000). It contained 70 different categories with 20 images per category yielding a total of 1400 binary images; each image contained a single shape, which we represented as a point set. We tested the LBO-Shape density technique on pose normalized and non-pose normalized versions of this dataset. The goal was to see if pose normalization had a significant effect on the classification performance. For the pose normalized case, shapes within each category were rigidly aligned to a category reference shape. The Brown dataset is a silhouette database from Sebastian et al. (2001) and contained nine classes with 11 shapes per class, yielding a total of 99 images. One interesting characteristic of this dataset is that many of the shapes have missing parts and occlusions. The first step was to construct an ϵ -neighborhood graph for each shape. Given that each shape was fit in a $[-10, 10] \times [-10, 10]$ domain, an ϵ value of 4 under the Euclidean metric yielded the best retrieval performance. Using these graphs, we computed the graph Laplacian for each shape (see

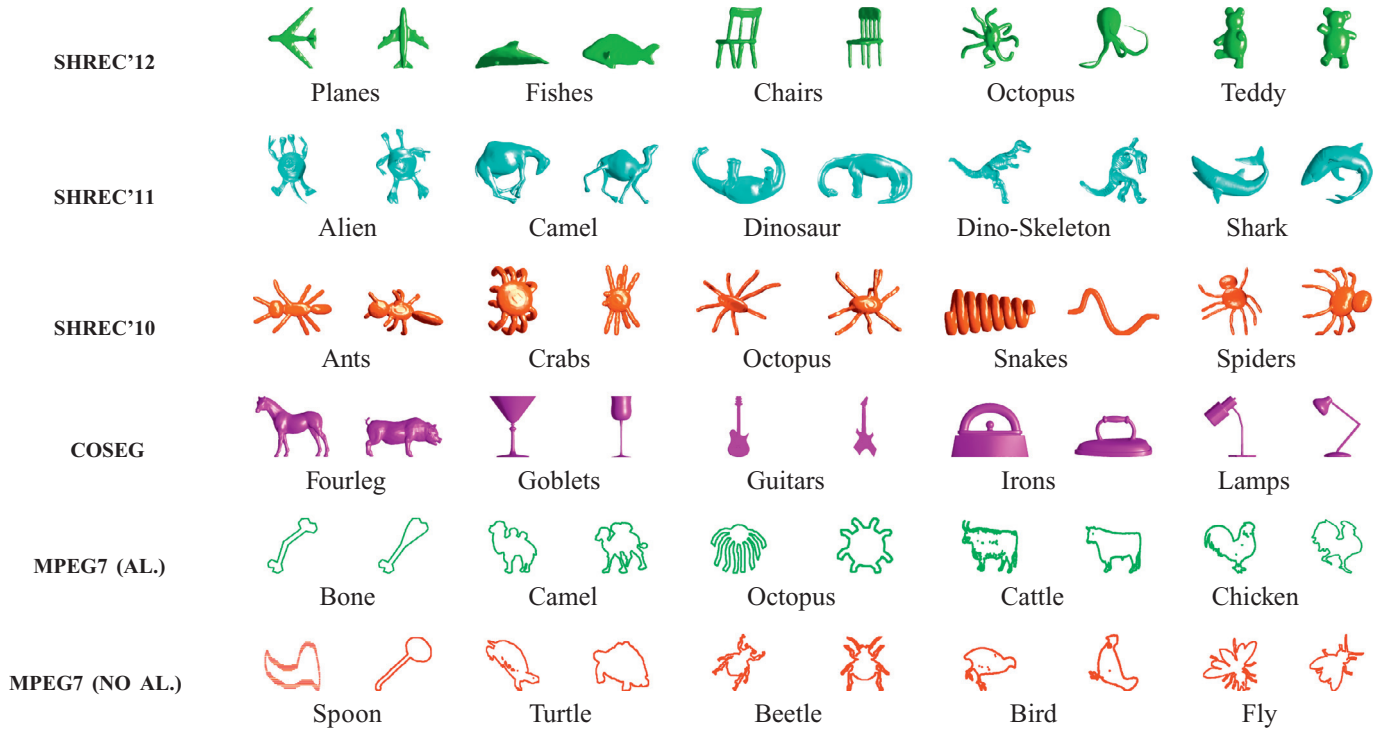


Fig. 5. Shapes from SHREC'12, SHREC'11, SHREC'10, Shape COSEG and MPEG-7 CE-Shape-1 part B datasets. Shapes from SHREC'15 and the Brown dataset are not shown. Notice that four of the categories on the SHREC'10 contain shape categories that are very similar to each other, such as ants, crabs, octopus and spiders. For the MPEG7 dataset the row (AL.) shows shapes that were aligned to a category reference shape. Whereas MPEG7 (NO AL.) shows shapes that were not aligned to a reference shape. See experimental results (Table 1) for retrieval performance with and without pose normalization.

Section 3.2) and obtained their corresponding eigenvectors. Note that the constant eigenvector corresponding to the zeroth eigenvalue is removed and not used in the classification.

Wavelet densities and their corresponding coefficients were estimated (see Section 3.5) on the triplets of eigenvectors for various wavelet and resolution combinations. A mean shape density was calculated for each category using the estimated coefficients as an input into Algorithm 1. For robustness testing on MPEG7, we randomly selected varying numbers of exemplars from each category to compute the mean density, beginning with all 20 shapes per category down to only 2 shapes per category. Finally, all 1400 shapes were treated as query shapes and the distances between these query shape densities and each of the class mean densities were computed—the class label of the mean density with minimum distance was assigned to the query shape. At this point, each mean shape density was being computed with a random subset of exemplars from a shape category. To be rigorous, we repeated the entire procedure 100 times to comprehend the effects of this

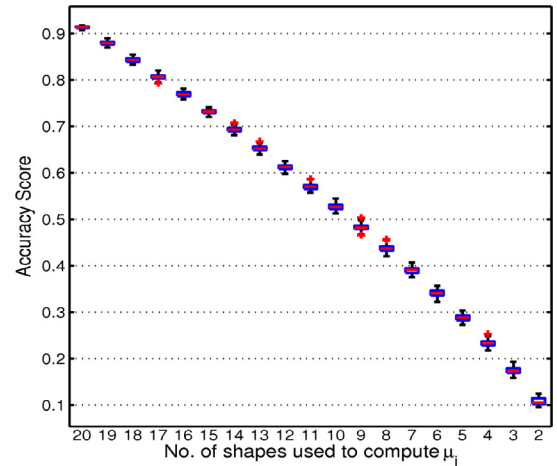


Fig. 6. MPEG-7 classification accuracies. The box plot shows the corresponding retrieval accuracy, over 100 random trials each, when reducing the number of shapes used to calculate the mean shape density of each category. We achieve a 91% accuracy with 20 shapes used to calculate the mean.

Algorithm 1 Numerical computation of Karcher mean on manifold M . For the present context $M = S^m$, m -dimensional unit hypersphere, the Exp and Log maps are defined in (19) and (20), respectively. κ is a small step size parameter.

Input: $\rho_1, \rho_2, \dots, \rho_m \in M$

Output: $\mu \in M$

Let $\mu^0 = \rho_1$

While $\|\gamma^\tau - \gamma^{\tau-1}\| > \epsilon$

$$\gamma^\tau = \frac{\kappa}{m} \sum_{i=1}^m \text{Log}_{\mu^{\tau-1}}(\rho_i)$$

$$\mu^\tau = \text{Exp}_{\mu^{\tau-1}}(\gamma^\tau)$$

random selection on our recognition performance. The box plot in Fig. 6 along with Table 1 summarizes our classification results. The box plot shows that the retrieval performance gracefully decreases as the number of shapes used to calculate the mean are reduced.

The methods (Felzenszwalb and Schwartz, 2007; McNeill and Vijayakumar, 2006) are based on hierarchical representations and have reported recognition rates above 85% on the MPEG-7 data set. In their approaches, shapes are represented by their boundary outlines and often times use less than 200 points for the shapes. These methods have the limitation of extracting oriented, boundary curves which can be a troublesome pre-processing procedure—they subsequently lose the descriptive power afforded by allowing arbitrary shape topologies and unconstrained point set cardi-

Table 1

MPEG-7 CE-Shape-1 part B 2D classification results based on the LBO-Shape density technique. The table shows the Bull's eye score (BES) and accuracy results obtained using different combinations of wavelets and resolution levels (j_0) for the pose normalized and non-pose normalized dataset. The highest BES and accuracy are achieved with a DB4 wavelet at resolution level 3. Similar retrieval accuracies with and without pose normalization are a direct result of the isometric invariance property of the LBO. For the Brown dataset we score 96% with $j_0 = 3$ and the (1,2,5) eigenvector combination. We are outperformed by the competing method only slightly. See experimental section for detail explanation of difference in performance.

Wavelet	j_0	Pose norm.		Non-pose norm.	
		BES	Acc.	BES	Acc.
Haar	3	0.914	0.845	0.913	0.822
DB2	3	0.950	0.913	0.949	0.893
DB4	3	0.951	0.913	0.946	0.890
BROWN					
GL		Furuya and Ohbuchi (2014)			
DB4, $j_0 = 2$, (1,2,5)	0.909	1			
DB4, $j_0 = 3$, (1,2,5)	0.960				
DB4, $j_0 = 3$, (1,2,5)	0.950				

nalities. More recently, graph transduction approaches (Egozi et al., 2010; Y et al., 2007) have been applied in conjunction with any existing similarity measure, yielding accuracy scores in the 92 – 99.99% range on MPEG-7; however, these approaches still rely on curve representations.

Wavelet shape densities have also been used by the authors in Peter et al. (2008); their method differs in that they directly incorporate non-rigid alignment via “sliding” of the wavelet coefficients and also do not compute mean densities for each of the categories. One must note that the accuracies for the competing algorithms (Egozi et al., 2010; Felzenszwalb and Schwartz, 2007; McNeill and Vijayakumar, 2006; Yang et al., 2007) were computed using the bull's eye criterion (Latecki et al., 2000), which is a more forgiving accuracy measure compared to our strict minimum distance classification. We perform well on the Brown dataset as shown in Table 1 with a score of 96% with the (1,2,5) combination. Confusion matrices of the 2D classification are shown in Fig. 7. It is very evident that our classification scheme works well due to the light colors along the main diagonal. Based on the results and comparison with competing methods, the LBO-Shape density technique has been proven to be a viable 2D shape matching algorithm that is on par with the state-of-the-art, while maintaining the desirable qualities of requiring less pre-processing, straightforward usage and computational efficiency.

4.2. 3D shape classification results

The 3D shape retrieval framework was evaluated on the five benchmark datasets previously mentioned, using the Nearest Neighbor (NN) metric. A key point to note about all of these experiments is that we do a model-to-mean classification and the competing methods do a model-to-model classification. We consider the NN score on these datasets as their classification accuracy. Also, in our framework we only have a single representative mean for each class compared to the other methods that utilize all the possible shapes in each class for comparisons. This makes classification easier in a model-to-model scheme because there is a higher chance of finding a similar shape. We perform very well on all datasets considering our disadvantage to the other methods. This model-to-mean classification scheme can be extended into a retrieval framework which would mean that we retrieve the most relevant samples in the majority of cases.

The SHREC'15 dataset (Lian et al., 2015) contains 1200 non-rigid watertight models separated into 50 classes with 24 shapes each.

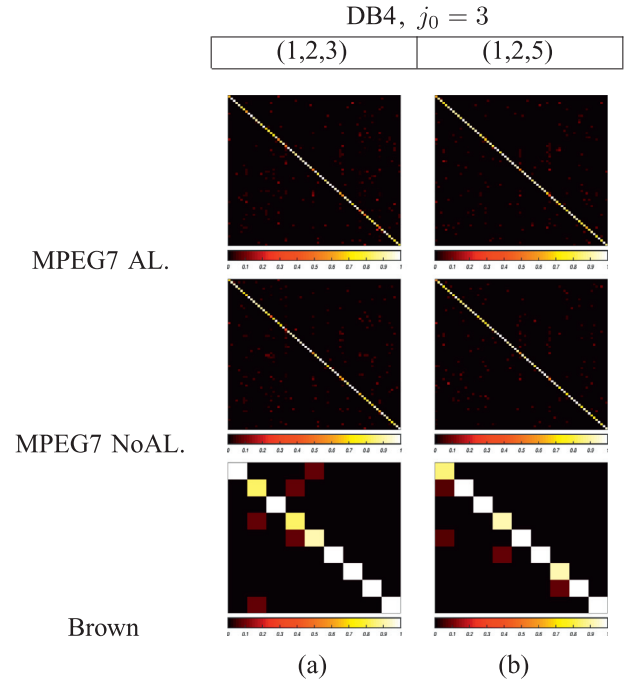


Fig. 7. The confusion matrices of the respective 2D datasets are shown in the rows of the table and the columns (a) and (b) denote the eigenvector triplets used, respectively. All results use a Daubechies 4 (DB4) basis function at resolution level $j_0 = 3$, note that the graph Laplacian is used for 2D shapes. The results show that our method performs well with most of the queries matched correctly as indicated by lighter colors along the main diagonals. MPEG AL and MPEG NoAL indicate results with and without pose normalization, respectively. The Brown dataset is not pose normalized. Our methods performs well even in the absence of pose normalization.

The SHREC'12 Track: Sketch-Based 3D Shape Retrieval dataset (Li et al., 2012), as applied in Khoury et al. (2012) contained 13 categories with 10 shapes per category in its basic version. Only the 3D models were selected as they provided non-similar categories such as ants, planes, and sunglasses. The SHREC'11 Track: Shape Retrieval on Non-rigid 3D Watertight Meshes dataset (Lian et al., 2011) contained 30 categories of watertight meshes containing 20 shapes each. The database contains a set of models which have similar overall appearances but belong to various categories because they are different in the details of local regions and topolog-

Table 2

Performance evaluation on SHREC'15 (a) SHREC'12 (b), SHREC'11 (c), SHREC'10 (d). The term GL refers to the graph Laplacian; CT refers to the cotangent approximation of the LBO; DB4 refers to the Daubechies 4 scaling basis function; j_0 refers to the resolution level; and the numbers proceeding this represent the particular triplet of eigenvectors chosen. On SHREC'12 we outperform the competing method (Khouri et al., 2012). On the remaining datasets we are only marginally outperformed by the competing methods. One reason for this difference in performance is due to the fact that we only compare to one shape (mean shape) to get a query label. This is different to the competing methods that compare to all the shapes in the class to make a classification. Notice the difference in performance between the different combinations of eigenvectors; overall the (1,2,5) combination performs better than the (1,2,3) combination. The higher the resolution level the better the performance in most cases. However, this is not necessarily the case when using the cotangent approximation of the LBO. The resolution level does not affect performance as much as the graph Laplacian. This is because the graph Laplacian is sensitive to the ϵ value used during the graph construction. The cotangent approximation does not have this free parameter.

	SHREC'10 (a)			SHREC'11 (b)			SHREC'12 (c)			SHREC'15 (d)		
	GL	CT	Ohbuchi et al. (2008)	GL	CT	Smeets et al. (2009)	GL	CT	Khouri et al. (2012)	GL	CT	Furuya and Ohbuchi (2014)
DB4, $j_0 = 2$, (1,2,5)	0.795	0.905	0.985	0.812	0.938	1	0.915	0.954	0.810	0.682	0.813	1
DB4, $j_0 = 3$, (1,2,5)	0.970	0.970		0.950	0.977		0.946	0.954		0.952	0.910	
DB4, $j_0 = 3$, (1,2,3)	0.953	0.980		0.948	0.948		0.946	0.946		0.943	0.882	

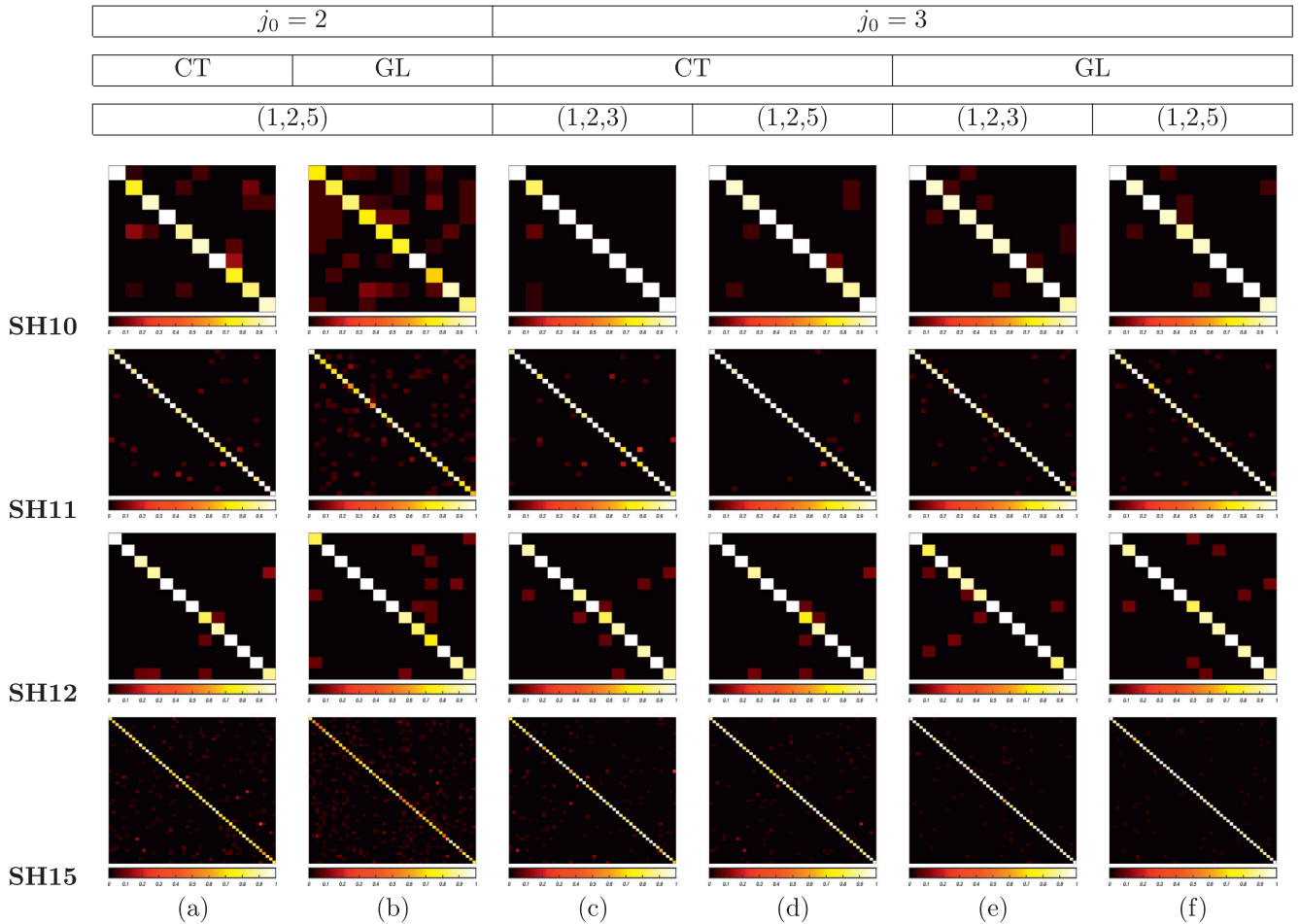


Fig. 8. The confusion matrices of the respective 3D datasets are shown in the rows of the table. CT represents the cotangent approximation to the LBO and GL represents the graph Laplacian. All results use a Daubechies 4 basis function with different resolution levels j_0 . The first two columns (a)-(b) show the results using $j_0 = 2$ with eigenvector triplet (1,2,5). Columns (c)-(e) show the results using $j_0 = 3$, note that results are presented for two eigenvector triplets (1,2,3) and (1,2,5). The results show that our method performs well at higher resolution levels with most of the queries matched correctly as indicated by lighter colors along the main diagonals.

ical structures. The SHREC'10 Track: Non-rigid 3D Shape Retrieval dataset (Lian et al., 2010) has 10 categories with 20 shapes per category and contains four visually similar categories: ants, crabs, spiders and octopi.

Our specific approach for 3D shape classification was to obtain the low order eigenvectors (1,2,5) of each shape's cotangent approximation to the LBO and graph Laplacian (after removal of the constant eigenfunction). The remaining steps are identical to those

for 2D matching, first the coefficients of the wavelet density were calculated on the triplet of eigenvectors using a scaling function at different resolution levels (see Table 2; wavelet functions are not used in this investigation) and resolution level j_0 . Then, these coefficients were used to compute the Karcher mean for each category, and form the mean index. Finally, shapes were classified using the minimum distance to the mean index. We used the Daubechies 4 scaling basis function in our experiments.

Across all of the datasets we achieve above 90% classification rate. This is visually depicted in Fig. 8, where the confusion matrices of the 3D datasets at different resolution levels are shown. The classification is very accurate as indicated by the lighter colors along the main diagonals. Note that our classification procedure is very different from the competing methods. We compare the query shape to a single mean shape versus the competing methods that compare to all the shapes from the class, hence they will have a greater likelihood of choosing the correct class. This is an important distinction to pay attention to because if a retrieval scheme is created based on this model-to-mean paradigm, we would attain very high retrieval scores.

In Table 2 (a) on the SHREC'15 dataset the graph laplacian (GL) performs better than the cotangent approximation (CTA) using the (1,2,5) combination with a score of 95%. For the SHREC'12 database, Table 2 (b) details our performance against the indexed heat curve approach in Khoury et al. (2012) where we outperform the competing method. The CTA outperformed GL with the (1,2,5) combination. An interesting point to note is that for the CTA the score for a resolution level of 2 is the same for resolution level 3. This is due to the small number of shapes in this dataset, and the stability of the CTA signature versus the GL that has a free parameter ϵ from the graph construction. Table 2 (c) shows the results for SHREC'11, once again the CTA outperforms the GL with a score of 97%. Recall that SHREC'11 has 600 shapes so we can see that the difference in scores between the resolution levels is not that significant for the CTA. For the GL there is a significant difference in scores between the resolution levels, which suggests that the CTA is a more stable 3D signature. For SHREC'10 in Table 2 (d) we see that the CTA gets the highest score of 98%, but with the (1,2,3) combination. The highest score for the (1,2,5) combination is 97%—it is only marginally higher, which suggests that the specific combination is not that significant.

5. Conclusion

We have presented strong supporting theoretical and empirical evidence that the LBO-Shape density technique for 2D/3D shape matching is indeed novel and robust in performance. Through the low order eigenvectors of different discretizations of the Laplace–Beltrami operator we are able to capture useful shape information. By computing 3D square-root wavelet densities on triplets of these eigenvectors, we represent shapes as density functions; in the process, shapes are furnished with a manifold structure that allows them to be identifiable with a point on a unit hypersphere. Using the hypersphere's noncomplex geometry, the Karcher means of these shape densities are intrinsically estimated on the manifold and serve as the new index used for retrieval.

In 2D, through the estimation of wavelet densities on the eigenspace spanned by three eigenvectors of the LBO (graph Laplacian approximation) we purge ourselves of point set cardinality constraints and topological restrictions, curb the required pre-processing, and adopt a computationally efficient similarity metric which is beneficial for large querying applications, while supporting sparse representations. Wavelet density estimation was performed for 3D shape retrieval using the cotangent approximation of the LBO and the graph Laplacian—harnessing the isometry invariance and intrinsic geometry encoding capabilities of its eigenvectors.

Our approach is also unique in that essentially the same identical process supports both 2D and 3D matching. The extensive experimental evaluations validated the utility of our approach, exhibiting very favorable competitive performance. The adoption of a model-to-mean paradigm versus the traditional model-to-model approach is well justified through the quantitative results. In the future, we intend to investigate the usefulness of different dis-

cretization schemes of the LBO, leverage the sparse representation capabilities of wavelets, and extend lessons learned to other areas of shape analysis like registration and segmentation.

Acknowledgments

The authors acknowledge partial support from NSF grant No. 1263011 and No. 1560345. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

- Amari, S.-I., Nagaoka, H., 2001. *Methods of Information Geometry*. American Mathematical Society.
- Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15, 1373–1396.
- Bimbo, A.D., Pala, P., 2006. Content-based retrieval of 3D models. *ACM Transactions on Multimedia Computing, Communications, and Applications* 2 (1), 20–43.
- Bronstein, A.M., Bronstein, M.M., Guibas, L.J., Ovsjanikov, M., 2011. Shape Google: geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics* 30, 1–20.
- Bronstein, M.M., Kokkinos, I., 2010. Scale-invariant heat kernel signatures for non-rigid shape recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*.
- Caelli, T., Kosinov, S., 2004. An eigenspace projection clustering method for inexact graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 515–519.
- Cao, Y., Zhang, Z., C., I., Dryden, I., Wang, S., 2011. 2D nonrigid partial shape matching using MCMC and contour subdivision. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2345–2352.
- Chen, T., Vemuri, B., Rangarajan, A., Eisenschenck, S.J., 2010. Group-wise point-set registration using a novel cdf-based Havrda–Charvát divergence. *International Journal of Computer Vision* 86, 111–124.
- Chui, H., Rangarajan, A., 2000. A new algorithm for non-rigid point matching. In: *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2. IEEE Press, pp. 44–51.
- Donoho, D., Johnstone, I., Kerkycharian, G., Picard, D., 1996. Density estimation by wavelet thresholding. *The Annals of Statistics* 24 (2), 508–539.
- Dziuk, G., 1988. Finite elements for the Beltrami operator on arbitrary surfaces. In: *Partial Differential Equations and Calculus of Variations*, pp. 142–155.
- Egozi, A., Keller, Y., Guterman, H., 2010. Improving shape retrieval by spectral matching and meta similarity. *IEEE Transactions on Image Processing* 19 (5), 1319–1327.
- Fan, C.R., 1997. *Spectral Graph Theory*. American Mathematical Society.
- Felzenszwalb, P.F., Schwartz, J.D., 2007. Hierarchical matching of deformable shapes. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8.
- Furuya, T., Ohbuchi, R., 2014. Fusing multiple features for shape-based 3D model retrieval. In: *British Machine Vision Conference*.
- Grünwald, P., 2005. A tutorial introduction to the minimum description length principle. *Advances in Minimum Description Length: Theory and Applications*. MIT Press.
- Isaac, C., 1984. *Eigenvalues in Riemannian Geometry*, vol. 115, second ed. Academic Press Professional, Inc.
- Isaacs, J., Roberts, R.G., 2011. Metrics of the Laplace–Beltrami eigenfunctions for 2D shape matching. In: *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3347–3352.
- Jian, B., Vemuri, B.C., 2011. Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis Machine Intelligence* 1633–1645.
- Jones, P.W., Maggioni, M., Schul, R., 2008. Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels. *Proceedings of the National Academy of Sciences* 105 (6), 1803–1808.
- Karcher, H., 1977. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics* 30 (5), 509–541.
- Khoury, R.G., Vandeborje, J.-P., Daoudi, M., 2012. Indexed heat curves for 3D-model retrieval. In: *IEEE International Conference on Pattern Recognition*, pp. 1964–1967.
- Latecki, L.J., Lakämper, R., Eckhardt, U., 2000. Shape descriptors for non-rigid shapes with a single closed contour. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 424–429.
- Levina, E., Bickel, P., 2001. The earth mover's distance is the Mallows distance: Some insights from statistics. In: *IEEE International Conference on Computer Vision*, vol. 2, pp. 251–256.
- Levy, B., 2006. Laplace–Beltrami eigenfunctions towards an algorithm that “understands” geometry. In: *IEEE International Conference on Shape Modelling and Applications*, p. 13.
- Li, B., Schreck, T., Godil, B., 2012. SHREC'12 track: Sketch-based 3D shape retrieval. *Eurographics Workshop on 3D Object Retrieval*.
- Li, C., Hamza, A.B., 2013. A multiresolution descriptor for deformable 3D shape retrieval. *The Visual Computer* 29 (6–8), 513–524.

- Lian, Z., Godil, A., Bustos, B., Daoudi, M., Hermans, J., Kawamura, S., Kurita, Y., Lavoué, G., Nguyen, H.V., Ohbuchi, R., Ohkita, Y., Ohishi, Y., Porikli, F., Reuter, M., Sipiran, I., Smeets, D., Suetens, P., Tabia, H., Vandermeulen, D., 2011. SHREC'11 Track: shape retrieval on non-rigid 3D watertight meshes. In: Eurographics Workshop on 3D Object Retrieval, pp. 79–88.
- Lian, Z., Godil, A., Fabry, T., Furuya, T., Hermans, J., Ohbuchi, R., Shu, C., Smeets, D., Suetens, P., Vandermeulen, D., Wuhler, S., 2010. SHREC'10 Track: non-rigid 3D shape retrieval. In: Eurographics Workshop on 3D Object Retrieval, pp. 101–108.
- Lian, Z., Zhang, J., Choi, S., ElNaghy, H., El-Sana, J., Furuya, T., Giachetti, A., Guler, R.A., Lai, L., Li, C., Li, H., Limberger, F.A., Martin, R., Nakanishi, R.U., Neto, A.P., Nonato, L.G., Ohbuchi, R., Pevzner, K., Pickup, D., Rosin, P., Sharf, A., Sun, L., Sun, X., Tari, S., Unal, G., Wilson, R.C., 2015. Non-rigid 3D shape retrieval. Eurographics Workshop on 3D Object Retrieval.
- Liu, Y., Prabhakaran, B., Guo, X., 2012. Point-based manifold harmonics. IEEE Transactions on Visualization and Computer Graphics 18 (10), 1693–1703.
- McNeill, G., Vijayakumar, S., 2006. Hierarchical procrustes matching for shape retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 885–894.
- Mingqiang, Y., Kidiyo, K., Joseph, R., 2008. A survey of shape feature extraction techniques. In: Yin, P.-Y. (Ed.), Pattern Recognition, pp. 43–90.
- Moyou, M., Ihou, K.E., Peter, A.M., 2014. LBO Shape Densities: efficient 3D shape retrieval using wavelet densities. In: IEEE International Conference on Pattern Recognition.
- Moyou, M., Peter, A.M., 2012. Shape analysis on the hypersphere of wavelet densities. In: IEEE International Conference on Pattern Recognition, pp. 2091–2094.
- Ohbuchi, R., Osada, K., Furuya, T., Banno, T., 2008. Salient local visual features for shape-based 3D model retrieval. In: IEEE International Conference on Shape Modeling and Applications, pp. 93–102.
- Patané, G., 2013. wFEM heat kernel: discretization and applications to shape analysis and retrieval. Computer Aided Geometric Design 30 (3), 276–295.
- Pennec, X., 2006. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. Journal of Mathematical Imaging and Vision 25 (1), 127–154.
- Peter, A.M., Rangarajan, A., 2008. Maximum likelihood wavelet density estimation with applications to image and shape matching. IEEE Transactions on Image Processing 17 (4), 458–468.
- Peter, A.M., Rangarajan, A., 2009. Information geometry for landmark shape analysis: unifying shape representation and deformation. IEEE Transactions on Pattern Analysis Machine Intelligence 31 (2), 337–350.
- Peter, A.M., Rangarajan, A., Ho, J., 2008. Shape L'Âne Rouge: sliding wavelets for indexing and retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8.
- Pinkall, U., Juni, S.D., Polthier, K., 1993. Computing discrete minimal surfaces and their conjugates. Experimental Mathematics 2, 15–36.
- Rangarajan, A., Chui, H., Bookstein, F., 1997. The softassign Procrustes matching algorithm. In: Information Processing in Medical Imaging. Springer, pp. 29–42.
- Reuter, M., Biasotti, S., Giorgi, D., Patané, G., Spagnuolo, M., 2009. Discrete Laplace–Beltrami operators for shape analysis and segmentation. Computer Graphics 33 (3), 381–390.
- Reuter, M., Wolter, F.-E., Niklas, P., 2006. Laplace–Beltrami spectra as ‘Shape-DNA’ of surfaces and solids. Computer-Aided Design 38 (4), 342–366.
- Rosenberg, S., 1997. The Laplacian on a Riemannian Manifold: An Introduction to Analysis on Manifolds. London Mathematical Society Student Texts. Cambridge University Press.
- Rubner, Y., Tomasi, C., Guibas, L.J., 2000. The earth mover’s distance as a metric for image retrieval. International Journal of Computer Vision 40 (2), 99–121.
- Ruggeri, M.R., Patané, G., Spagnuolo, M., Saupe, D., 2010. Spectral-driven isometry-invariant matching of 3D shapes. International Journal of Computer Vision 89, 248–265.
- Rustamov, R.M., 2007. Laplace–Beltrami eigenfunctions for deformation invariant shape representation. Eurographics Symposium on Geometry processing.
- Sang, H.P., Kyoung, M.L., Sang, U.L., 2000. A line feature matching technique based on an eigenvector approach. Computer Vision and Image Understanding 77 (3), 263–283.
- Sebastian, T.B., Klein, P.N., Kimia, B.B., 2001. Recognition of shapes by editing shock graphs. In: IEEE International Conference on Computer Vision, vol. 1, pp. 755–762.
- Siddiqi, K., Shokoufandeh, A., Dickinson, S.J., Zucker, S.W., 1998. Shock graphs and shape matching. In: IEEE International Conference on Computer Vision, pp. 222–229.
- Smeets, D., Fabry, T., Hermans, J., Vandermeulen, D., Suetens, P., 2009. Isometric deformation modelling for object recognition. In: Computer Analysis of Images and Patterns, pp. 757–765.
- Srivastava, A., Joshi, S.H., Washington, M., Xiuwen, L., 2005. Statistical shape analysis: clustering, learning, and testing. IEEE Transactions on Pattern Analysis Machine Intelligence 27 (4), 590–602.
- Srivastava, A., Jermyn, I., Joshi, S., 2007. Riemannian analysis of probability density functions with applications in vision. In: IEEE Computer Vision and Pattern Recognition, pp. 1–8.
- Sun, J., Ovsjanikov, M., Guibas, L., 2009. A concise and provably informative multi-scale signature based on heat diffusion. In: Symposium on Geometry Processing, pp. 1383–1392.
- Tangelder, J.W., Veltkamp, R.C., 2008. A survey of content based 3D shape retrieval methods. Multimedia Tools and Applications 39 (3), 441–471.
- Thakoor, N., Gao, J., Jung, S., 2007. Hidden markov model-based weighted likelihood discriminant for 2-d shape classification. IEEE Transactions on Image Processing 16 (11), 2707–2719.
- Umeyama, S., 1988. An eigendecomposition approach to weighted graph matching problems. IEEE Transactions on Pattern Analysis Machine Intelligence 10 (5), 695–703.
- Wardetzky, M., Mathur, S., Kälberer, F., Grinspun, E., 2007. Discrete Laplace operators: no free lunch. In: Symposium on Geometry processing, pp. 33–37.
- White, D., Wilson, R., 2007. Spectral generative models for graphs. In: International Conference on Image Analysis and Processing, pp. 35–42.
- Xu, G., 2004. Discrete Laplace–Beltrami operators and their convergence. Computer Aided Geometric Design 21 (8), 767–784.
- Yang, Y., Lin, H., Zhang, Y., 2007. Content-based 3D model retrieval: a survey. IEEE Transactions on Systems, Man, and Cybernetics 37 (6), 1081–1098.
- Zeng, W., Guo, R., Luo, F., Gu, X., 2012. Discrete heat kernel determines discrete Riemannian metric. Graphical Models 74 (4), 121–129.