Shape-Based Retrieval of Articulated 3D Models Using Spectral Embedding

Varun Jain and Hao Zhang

GrUVi Lab, School of Computing Sciences,
Simon Fraser University, Burnaby, British Columbia, Canada
{vjain, haoz}@cs.sfu.ca

Abstract. We present an approach for robust shape retrieval from databases containing articulated 3D shapes. We represent each shape by the eigenvectors of an appropriately defined affinity matrix, obtaining a spectral embedding. Retrieval is then performed on these embeddings using global shape descriptors. Transformation into the spectral domain normalizes the shapes against articulation (bending), rigid-body transformations, and uniform scaling. Experimentally, we show absolute improvement in retrieval performance when conventional shape descriptors are used in the spectral domain on the McGill database of articulated 3D shapes. We also propose a simple eigenvalue-based descriptor, which is easily computed and performs comparably against the best known shape descriptors applied to the original shapes.

1 Introduction

In recent years, there has been a tremendous advance in 3D model acquisition technology and a large number of 3D models have become available on the web or through other means. The problem of indexing and retrieval of 3D shapes [1] has become as important, both in practice and in terms of research interests, as that of indexing and retrieval of image or textual data. Formally, given a database of 3D shapes represented in the form of triangle meshes[1], and a query shape, a shape retrieval algorithm seeks to return shapes, ordered by decreasing visual similarity to the query shape, from the database that belong to the same class as the query, where the classification is done by human.

Since the process of object recognition by human has not been completely understood, we are still incapable of proving theoretically that one particular shape retrieval algorithm is the best. In practice, several benchmark data sets and their associated performance evaluations [1–3] are available to empirically measure the quality of existing shape retrieval algorithms. The most comprehensive comparative study of retrieval algorithms for 3D shapes to date is due to Shilane et al. [1], based on the now well-known Princeton shape benchmark.

[1] A liberal use of the term mesh is adopted: the mesh can be non-manifold, open or closed, having disconnected parts, or a collection of disjoint soup of triangles.

A variety of retrieval algorithms have been proposed [4]. Typically, each shape is characterized by a shape descriptor. An appropriately defined similarity distance between the descriptors sorts the retrieved models. Commonly used quality criteria for shape descriptors include invariance to rigid-body transformations, scaling, bending and moderate stretching, robustness against noise and data degeneracies, and storage and computational costs. The discriminative power of a shape descriptor and its similarity distance is most often judged by plotting the precision-recall (PR) curve [1] generated from a benchmark database.

Most state-of-the-art descriptors, including the twelve compared by Shilane et al. [1] on the Princeton benchmark, are designed to be invariant to only rigid-body transformations and uniform scaling. Hence, it is no surprise that they do not perform well when applied to shapes having non-rigid transformations such as bending or stretching, which are obviously harder to handle due to their non-linearity and increased degrees of freedom. In this paper, we propose a technique to render a descriptor invariant to bending, hence enhancing its performance over databases that contain articulated shapes. Our experiments will thus be conducted primarily on the McGill database of articulated shapes [3].

Given a shape represented as a triangle mesh, we first apply pre-processing to convert it into a connected weighted graph. Shortest graph distances between pairs of nodes, mimicking geodesic distances over the mesh surface, provide an intrinsic characterization of the shape structure. We filter these distances appropriately to remove the effect of scaling and then compute a spectral embedding of the shape in a low-dimensional space, where we attain invariance to bending. The spectral embeddings are given by the eigenvectors, properly scaled, of the matrix of filtered distances. The corresponding eigenvalues can be used to obtain a simple shape descriptor that works quite well on the McGill database.

Alternatively, any existing 3D shape descriptors can be applied to spectral embeddings in 3D, which would result in absolute performance improvements in the PR curves on the McGill database. In this paper, we demonstrate this for the spherical harmonic shape (SHD) descriptor of Kazhdan et al. [5] and the light field descriptor (LFD) of Chen et al. [2], two of the best-performing descriptors from the Princeton benchmark test [1]. Finally, it is worth noting that with the aid of sub-sampling and interpolation via Nyström approximation [6], the spectral embeddings are quite efficient to compute.

The rest of the paper is organized as follows. After briefly discussing previous work in Section 2, we describe efficient construction of the bending-invariant spectral embeddings from a given mesh, possibly with disconnected components and other degeneracies, in Section 3. In Section 4, we give a comparative study between various shape descriptors, including those derived from spectral embeddings, for shape retrieval. Experimental results and discussions are given in Section 5. Finally, we conclude in Section 6 and suggest possible future work.

2 Previous work

It is quite conceivable that a great deal of prior knowledge is incorporated into the process of human object recognition and classification, perhaps with subpart

matching playing an important role. In this paper however, we focus on purely shape-based approaches using global shape descriptors [4]. At a high level, a 3D shape retrieval algorithm either works on the 3D models directly, e.g., [5, 7], or relies on a set of projected images [2] taken from different views. Let us call these the object-space and the image-space approaches, respectively. The latter, e.g., the LFDs of Chen et al. [2], has a more intuitive appeal to visual perception and thus often result in better benchmark results for retrieval [1], but at the expense of much higher computational cost.

Many object-space shape descriptors construct one or a collection of spherical functions, capturing the geometric information in a 3D shape *extrinsically* [1]. These spherical functions represent the distribution of one or more quantities, e.g., distance from points on the shape to the center of mass [8], curvatures [9], surface normals [10], etc. The bins are typically parameterized by the sphere radius and angles. The spherical functions are, in most cases, efficient to compute and robust to geometric and topological noise, but they may be sensitive to the choice of sphere center or the bin structures. To align the bins for two shapes properly, these approaches require pre-normalization with respect to translation, rotation, and uniform, e.g., [8, 10], or nonuniform scaling [11]. As an alternative, rotation-invariant measures computed from the spherical functions, e.g., energy norm at various spherical harmonic frequencies [5], can be utilized. However, non-rigid transformations are not handled by these approaches.

As a salient intrinsic geometric measure, surface curvature, as well as the principal curvature directions, have been used for shape characterization and retrieval [9]. These approaches are sensitive to noise and non-rigid transforms such as bending. Another intrinsic approach is the use of shape distributions [7], where a histogram of pairwise distances between the vertices of a mesh defines the shape descriptor. Other form of statistics, e.g. [12], can also be used and bending-invariance can obviously be achieved if geodesic distances are used in this context, but the discriminative power of the histograms is suspect.

The most common approach to handling shape articulation is via skeletal or other graph representation of the shapes, e.g., [13, 14], and then apply graph matching. The cost of extracting the skeletons can be high, e.g., when medial-axes are used [14], and the subsequent graph matching is often computationally expensive and the shape descriptor itself is sensitive to topological noise. Our approach also uses a graph-based intrinsic characterization of the shapes. The spectral embeddings automatically normalize the shapes against rigid-body transformations, uniform scaling, and bending, and they are fast to compute. The resulting shape descriptors provide a more intuitive way of characterizing shapes, compared to shape distribution [7]. In addition, the spectral approach is quite flexible and allows for different choices of graph edge weights and distance computations, rendering the approach more robust against topological noise.

The idea of using spectral embeddings for data analysis is not new and clustering [15, 16] and correspondence analysis [17–19] are the main applications. Past work that is most relevant to ours is the use of bending-invariant shape signatures by Elad and Kimmel [20]. They work on manifold meshes and compute

spectral embeddings using multidimensional scaling (MDS) based on geodesic distances. A more efficient version of MDS is adopted to approximate the true embeddings; this is different from Nyström approximation. They only tested shape retrieval on manifold, isometric shapes, e.g., models obtained by bending a small set of seed shapes. In practice, many 3D shapes are neither manifolds nor isometric to each other, thus a more robust approach, based on more general graphs and distance measures, and a more complete experiment, are called for.

3 Construction of spectral embeddings

Point correspondence between two images or extracted image features has been well studied in computer vision. Spectral technique is first applied to this problem by Umeyama [21], Scott and Longuet-Higgins [22], and Shapiro and Brady [18]. Since then, the use of spectral techniques for correspondence in 2D has received a great deal of attention [17]. In machine learning, spectral clustering [15] and its variants are well-known. However, the use of spectral methods for 3D geometry processing is relatively new. For example, Zhang and Liu [16] apply spectral embeddings to mesh segmentation, while Gotsman et al. [23] utilize the spectral properties of mesh Laplacians for spherical parameterization. Spectral analysis has also been applied to mesh compression [24]. To the best of our knowledge, the use of spectral embeddings for 3D shape retrieval has not been reported before. In this section, we describe the process of constructing spectral embeddings for a 3D mesh that can be subsequently used for shape retrieval.

3.1 Affinity matrix and spectral embedding

Given a 3D triangle mesh with n vertices, we form an $n \times n$ affinity matrix A such that the ij^{th} entry of A is the affinity between the i^{th} and the j^{th} mesh vertices. Several possible choices for the affinities are discussed in Section 3.4. A is then eigen-decomposed as, $A = V \Lambda V^T$, where Λ is a diagonal matrix with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ along the diagonal and $V = [\mathbf{v_1}|\ldots|\mathbf{v_n}]$ is the $n \times n$ matrix of the corresponding eigenvectors, with $\mathbf{v_1}, \ldots, \mathbf{v_n}$ the eigenvectors.

As the eigenvectors are of unit length, their entries may vary in scale with change of the mesh size n. We normalize this variation by scaling the eigenvectors by the square-root of the corresponding eigenvalues [19] and then consider only the first k scaled eigenvectors to give a k-dimensional *spectral embedding*:

$$\hat{V}_k = [\hat{\mathbf{v}}_\mathbf{1}|\ldots|\hat{\mathbf{v}}_\mathbf{k}], \text{where } \hat{\mathbf{v}}_\mathbf{1}, \ldots, \hat{\mathbf{v}}_\mathbf{k} \text{ are the first } k \text{ columns of } \hat{V} = V\Lambda^{\frac{1}{2}}.$$

Specifically, the i^{th} row of the $n \times k$ matrix $\hat{V}_k$ gives the k-dimensional coordinates of the i^{th} vertex of the mesh.

An advantage of using this particular framework for shape characterization is that if the affinities in the matrix A are invariant to a particular transformation, then the resultant embeddings will also be invariant to that transformation. This property can be exploited to construct bending-invariant embeddings if we note that the geodesic distance between two points on a mesh remains constant when the shape undergoes bending. We now explain the construction of bending-invariant spectral embeddings using approximate geodesic distances.

3.2 Bending-invariant spectral embedding

In order to achieve invariance to bending, we wish to define the affinities based on geodesic distances. However, conventional methods for geodesic estimation over a mesh depend largely on the mesh being connected. This limits the use of geodesic distances, as we have noticed that many shapes in all the well-known shape database [1, 3] have disconnected parts (a small number of shapes are simply triangle soups), in which case, the geodesic distance estimation would fail. We thus turn to a heuristic as a work-around.

Construction of structural graph: We use shortest graph distances over a mesh graph to approximate geodesic distances. This not only leads to simpler implementation, but also removes the constraint that the shape be defined using a connected manifold mesh. However, disconnected meshes are still not handled properly. To this end, we add extra edges to the mesh graph composed of its original vertices and edges while making sure that the structure of the shape remains largely unchanged. Specifically, given a 3D mesh M, let $G_M = (V, E_M)$ be its connectivity graph and $C_1, C_2, \ldots$ be its disconnected components. We construct a p-connected graph $G_p = (V, E_p)$ over the mesh vertices such that the graph faithfully represents the shape. This is done using Yang's algorithm [25] for constructing p-connected graph over point clouds in Euclidean space that locally minimizes edge lengths by computing and combining p Euclidean minimum spanning trees of the given point cloud. As shown in [25] and verified by our experiments, the resulting graph G_p approximates the structure of the shape well. With G_M and G_p in hand, the final *structural graph* is defined as

$$G = (V, E), \text{where } E = E_M \cup \{(i, j) \;\mid\; (i, j) \in E_p, i \in C_s, j \in C_t, s \neq t\}.$$

Clearly, G is connected and it includes all edges of G_M and only those edges of G_p that join two disconnected components in G_M. This helps better preserving the structure of the mesh. Once the structural graph G has been constructed, the geodesic distance between two vertices can be approximated by the shortest path length in G computed using Dijkstra's shortest path algorithm.

In our implementation, we restrict p, in Yang's algorithm, to be 1 or 2 since higher values of p may result in edges between far away (hence, unrelated) components. This is illustrated in Fig. 1 where we plot the average error (as a percentage of the bounding box diagonal or BBD) in the estimation of geodesic distances using the approach described above, against degeneracy levels in the mesh. To add degeneracy in a mesh we randomly select a number of faces and disconnect them from the mesh and jitter the position of their vertices. The noise level plotted on the horizontal axis indicates the number of such faces. The meshes on which the experiment was carried out all contain 350 faces. Clearly, increasing p would result in non-robustness of geodesic distance estimation.

Gaussian affinities: Now that we have a way to robustly estimate geodesic distances, we can define the affinity matrix A, which is given by a *Gaussian*: $A_{ij} = \exp\left(-d_{ij}^2/2\sigma^2\right)$, where d_{ij} is the approximate geodesic distance between the i^{th} and j^{th} vertices of the mesh, and σ is the Gaussian width. We simply

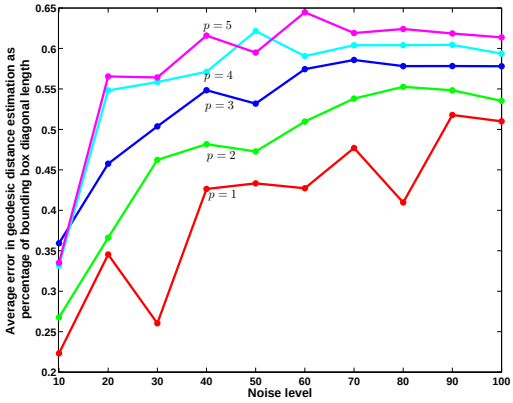


Fig. 1. Error (as % of BBD) in geodesic distance approximations with varying p.

Fig. 2. Spectral embeddings (bottom row) of some articulated 3D shapes (top row) from the McGill shape database. Note that normalizations have been carried out.

set $\sigma = \max_{(i,j)}\{d_{ij}\}$. Defining σ in this data-dependent manner renders the embedding invariant to uniform scaling. We observe experimentally that the embeddings are relatively stable with respect to σ as long as it is sufficiently large. As a consequence of setting σ to a large number, the row-sums of the matrix A become almost constant. It follows that the first eigenvector $\mathbf{v_1}$ of A is very close to being a constant vector. Hence, we exclude the first eigenvector and consider only a $(k-1)$-dimensional embedding defined by $\mathbf{v_2}, \ldots, \mathbf{v_k}$.

For all the 3D shape retrieval results obtained based on spectral embeddings in this paper, we represent every shape with a 3D spectral embedding given by the 2^{nd}, 3^{rd} and 4^{th} eigenvectors, scaled by the square-root of the corresponding eigenvalues, of the affinity matrix associated with the shape. The 3D embeddings of some articulated shapes from the McGill database are shown in Fig. 2.

3.3 Nyström approximation

The time complexity for constructing the full affinity matrix A for a mesh with n vertices is $O(n^2 \log n)$. Moreover, the eigen-decomposition of an $n \times n$ matrix takes $O(n^3)$ time; $O(kn^2)$ if only the first k eigenvectors are computed. This complexity does not affect the retrieval performance drastically, since the

spectral embeddings of all the shapes in the database can be precomputed. However, the query model needs to be processed at run-time. To speed things up, we use Nyström approximation [6] to efficiently approximate the eigenvectors of A. Nyström approximation is a sub-sampling technique that reduces the time complexity of affinity matrix construction and eigen-decomposition to $O(ln \log n + l^3)$, where l is the number of samples selected; typically, $l \ll n$. We adopt furthest point sampling, which at each step, chooses a sample which maximizes the minimum (approximated) geodesic distance from the new sample to the previously found samples; the first sample can be chosen randomly. Our extensive experiments confirm that for the purpose of shape retrieval, only 10 to 20 samples, from meshes with thousands of vertices, are sufficient.

3.4 Other affinity measures

Although geodesic-based affinities lead to bending invariance, it might cause adverse effects in some cases. For example, consider two chair models. Suppose that the arm-rest of one model is connected directly to its back-rest, however, on the other model, this connection is only through the seat. In the first case, the geodesic distance between a point on the arm-rest and a point on the back-rest is small, whereas in the second case this distance will be relatively large. Hence, the spectral embeddings of the two chairs could be radically different and the retrieval result will suffer. In general, geodesic distances are sensitive to topological noise in the shapes. If we define affinities based on Euclidean distances, the above problem would be resolved but the affinities can no longer be expected to be invariant (or even robust) to bending. Nevertheless, the current discussion reveals the flexibility of our approach, with the use of affinity matrices, in that they can be easily tuned to render the retrieval process invariant to a particular class of transformations depending on the database in question.

In Section 4, we compare retrieval results using different affinity measures. In addition to (approximated) geodesic-based affinities and affinities based on Euclidean distances, we also include *combined distance*, where a uniform combination of the above two measures is used. Since our target database is that of articulated shapes, it is not surprising that the geodesic-based affinities perform the best. Minor improvements over conventional shape descriptors can still be seen using other affinities, which strengthens our proposal of performing retrieval on spectral embeddings instead of on the original shapes.

4 Global descriptors for shape retrieval

We now present a comparative study of two global shape descriptors, the spherical harmonics descriptor [5] (SHD) and the light field descriptor [2] (LFD), in the context of shape retrieval. Both of these descriptors have been shown to give excellent shape retrieval results in the Princeton shape benchmark [1]. In fact, the light field descriptor is the best among all the descriptors compared in [1]. We evaluate the performance of the descriptors when they are applied to the

original meshes as compared to when they are applied to the spectral embeddings of the meshes. We use the McGill database of articulated 3D shape [3] for our experiments. We also present two simple descriptors, easily obtained from the spectral embeddings, that perform comparably to the other descriptors.

The McGill shape database contains 255 models in 10 categories: Ants, Crabs, Hands, Humans, Octopuses, Pliers, Snakes, Spectacles, Spiders, and Teddybears. There are 20 to 30 models per category. Some shapes from the database are shown in Fig. 2 and 4. We now explain the descriptors we compare.

1. **Light Field Descriptor (LFD) [2]**: represents the model using histograms of 2D images of the model captured from a number of positions, uniformly placed on a sphere. The distance between two models is the distance between the two descriptors minimized over all rotations between the two models, hence attaining robustness to rotations. The main idea of this descriptor is to define shape similarity based on the visual similarity of the two shapes.
2. **Spherical Harmonics Descriptor (SHD) [5]**: is a geometry based representation of the shape which is invariant to rotations. It is obtained by recording the variation of the shape using spherical harmonic coefficients computed over concentric spherical shells.
3. **Spectral Shape Descriptors**: The following are two shape descriptors that can be easily obtained from the spectral embeddings. The EVD descriptor consists of only six eigenvalues and performs comparably as and sometimes better than SHD. This shows the effectiveness of the affinity matrix and spectral embeddings in encoding essential shape information.
   (a) **Eigenvalue Descriptor (EVD)**: While the eigenvectors of the affinity matrix form the spectral embedding which is a normalized representation of the shape, the eigenvalues specify the variation of the shape along the axes given by the corresponding eigenvectors. Hence, as a simple shape descriptor, we use the square root of the first six eigenvalues of the affinity matrix. The reason for choosing only six eigenvalues is that the remaining terms in the spectrum are believed to encode high frequency shape information, which may render the descriptor too sensitive to shape noise. Also, the eigenvalues tend to decrease quickly, hence, only the largest eigenvalues shall encode significant shape information. With EVD, the distance between two meshes P and Q is given by the χ^2-distance between the square root of their first six eigenvalues:

$$Dist_{EVD}(P,Q) = \frac{1}{2}\sum_{i=1}^{6} \frac{[\ |\lambda_i^P|^{\frac{1}{2}} - |\lambda_i^Q|^{\frac{1}{2}}]^2}{|\lambda_i^P|^{\frac{1}{2}} + |\lambda_i^Q|^{\frac{1}{2}}}.$$

   It is worth noting however that the eigenvalues are affected by mesh sizes and there are shapes with different number of vertices in the shape database. Thus the eigenvalues of the original affinity matrices cannot be used for shape comparison directly. However, recall that we only compute a sampled affinity matrix, required by Nyström approximation. Thus

with the same number of samples taken on each shape, the eigenvalues of the sampled affinity matrices can be used as is.

(b) **Correspondence Cost Descriptor (CCD)**: The distance between two shapes in the CCD scheme is derived from the correspondence between the vertices of the two shapes. Given the respective k-dimensional spectral embeddings of two shapes P and Q in the form of an $n_P \times k$ matrix V_P and an $n_Q \times k$ matrix V_Q, the CCD distance given by:

$$Dist_{CCD}(P,Q) = \sum_{p \in P} \|V_P(p) - V_Q(match(p))\|,$$

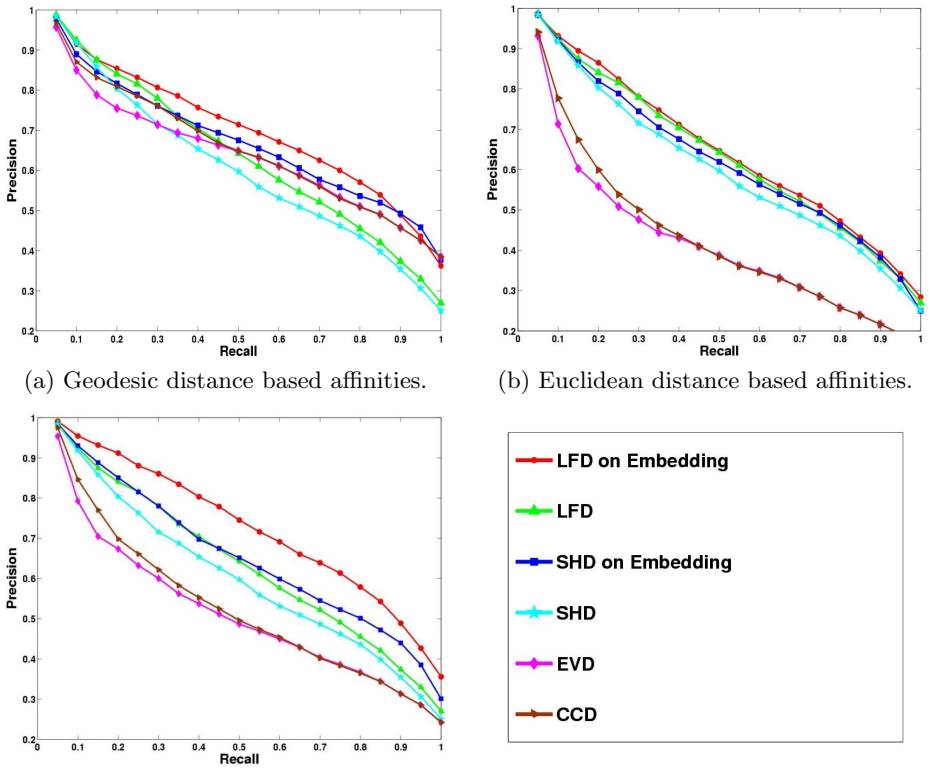
where $V_P(p)$ and $V_Q(q)$ are the p^{th} and q^{th} rows of V_P and V_Q, respectively, p represents a vertex of P, and $match()$ is some computed mapping between the vertices of P and the vertices of Q. This matching can be obtained using any correspondence algorithm, e.g., [18, 19]. We have chosen to compute correspondence using the spectral embeddings obtained from the previous step. The correspondence algorithm uses best matching based on Euclidean distance in the embedding space [18],

$$match(p) = \text{ argmin }_{q \in Q} \|V_P(p) - V_Q(q)\|.$$

The intuition behind defining such a similarity cost is that if two shapes are similar (though they may differ by a bending transformation), their spectral embeddings would be similar, hence the Euclidean distance between a point and its match will be small, resulting in a smaller value of $Dist_{CCD}(P,Q)$. However, note that the time complexity of finding the distance between two shapes in the CCD scheme is $O(n^2)$, where n is the number of vertices. This is extremely slow and is not feasible to apply for comparing the query model with a large number of models in a database. Hence, we use CCD in conjunction with EVD. We first use EVD to filter out all poor matches via thresholding. Only the top few matches obtained from EVD are further refined using CCD.

5 Experimental results

In this section, we present experimental results. First, we plot the precision-recall (PR) curves for the four descriptors, given in the previous section, when they are applied to the McGill database of articulated shapes, in Fig. 3. For (a), the approximate geodesic distances are used to construct the affinities. Clearly, the descriptors show significant improvements when applied to bending-invariant embeddings, compared to their spatial domain counterparts. In (b), we show the performance of the same set of descriptors. However, we construct embeddings based on Euclidean distances. Note that the performance of spectral descriptors degrades considerably. This is mainly because these are naive descriptors that rely on the ability of the affinity matrix to normalize proper transformations between the shapes. Since Euclidean distance based affinity matrix does not



(a) Geodesic distance based affinities. (b) Euclidean distance based affinities.

(c) Combined distance based affinities. Legends.

Fig. 3. Precision-recall (PC) plots for various global descriptors, derived from different distance measures, when applied to the McGill database of articulated shapes [3].

normalize the shapes against bending and the database in question is particularly that of articulated shapes, such poor performance is expected.

Fig. 3(c) shows the performance of the descriptors where the affinities are calculated using an average of geodesic and Euclidean distances. For LFD and SHD, both geodesic affinities and combined affinities give considerable improvements. Whereas, Euclidean affinities show only minor improvements since they fail to normalize the shapes against bending. EVD and CCD perform well only when the embeddings are normalized against bending for reasons mentioned above.

In terms of running times, the EVD's are quite efficient to compute due to subsampling. The time taken to compute the subsampled affinity matrix grows linearly with mesh size while the time required for computing the descriptor, an eigenvalue problem of size $k \times k$, is the same as we select $k = 10$ samples throughout. On an Intel Pentium M 1.7GHz machine with 1GB RAM, it takes between 1.4 and 2 milliseconds to compute the EVDs of meshes whose face counts range from 2,000 to 4,000. SHD and LFD computations are more expensive,

taking an average of about 2 and 2.3 milliseconds for meshes in that range of face counts, respectively.

Next, we show some visual results which emphasize the need for bending-invariant spectral embeddings in order to obtain more robust retrieval of articulated shapes. These results are shown in Fig. 4. The spectral embeddings used in these results are all constructed using geodesic distance based affinities. Fig. 4(a) shows the results of retrieving an ant shape from the database. Note the poor performance of SHD even when the amount of bending is moderate. Fig. 4(b) and (c) show results for querying the database with a human and a plier shape, respectively, that have a relatively larger amount of bending. As we can see, LFD performs rather poorly on the original shapes.

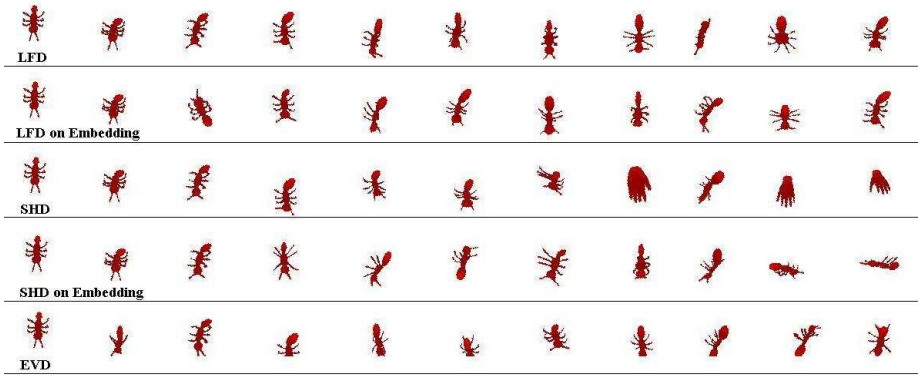
It is quite evident from Fig. 4 that shape descriptors applied to spectral embeddings show clear and consistent improvement over their spatial domain based counterparts. It is also interesting to note from Fig. 4 that EVD, our simple shape descriptor based on eigenvalues, appears to work the best. We have indeed observed that most, if not all, incorrect retrieval results using EVD are caused by having parts of a shape incorrectly connected in our construction of the structural graph. Recovering the correct shape information from a soup of triangles or sparsely and nonuniformly sampled points (which occur often in the shape databases) is not an easy problem, but any improvements in this regard will improve the performance of the EVD. Our current heuristic is quite primitive and we would like to look into this problem in our future work.

In Fig. 5, we show an image representation of the similarity matrix for all the shapes of the database. Here, a bright pixel represents greater similarity. The descriptor used here is LFD on spectral embedding instead of the original shape. The diagonal structure of the image matrix shows that similar shapes have greater similarity value.

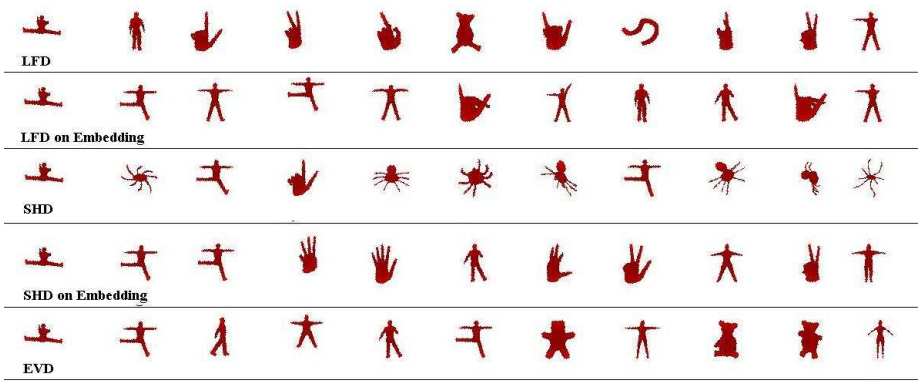
6 Conclusion and future work

In this paper, we consider the problem of shape-based retrieval of 3D models from a database and our focus is on articulated shapes. We present a method which renders conventional shape descriptors invariant to shape articulation, with the use of spectral embeddings derived from an appropriately defined affinity matrix. The affinity matrix encodes pairwise relations between the data points and invariance to a particular type of transformation can be achieved through a judicious choice of a distance measure. When conventional shape descriptors, e.g., LFD and SHD, are applied to spectral embeddings derived from approximated geodesic distances, on the McGill database of articulated 3D shapes, absolute improvements are achieved for shape retrieval. Robustness of affinity matrices is also shown, as minor improvements for the LFD and SHD descriptors can be observed, on the same database, even with Euclidean distance based affinities that are not invariant to bending.

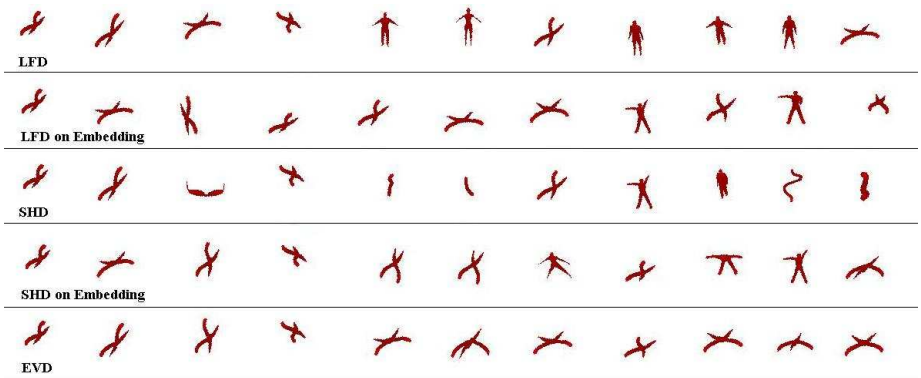
In the future, we would like to explore more ways to define affinities that are robust and/or invariant to other complex shape transformations such as non-uniform linear scaling and moderate stretching (note that allowing arbitrary



(a)



(b)



(c)

Fig. 4. Retrieval results using the McGill database of articulated 3D shapes [3]: First column in each row is the query shape. This is followed by the top ten matches retrieved using the shape descriptor as indicated.

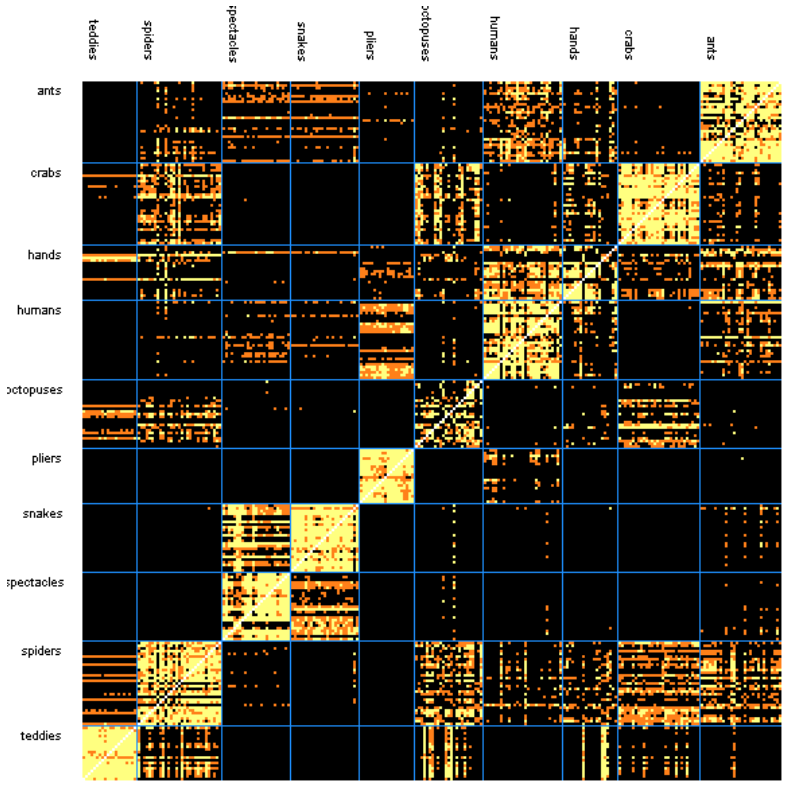


Fig. 5. Similarity matrix for shapes from the McGill database of articulated 3D shapes [3], computed using LFD on spectral embeddings.

stretching and bending would then only enable us to distinguish shapes having different topology). Another interesting study would be to find other shape descriptors based on spectral embeddings that can be used for retrieval; we have suggested two simple ones, EVD and CCD, in this paper and their performances are only about equivalent to that of the state of the art. Issues such as, the number of eigenvalues or eigenvectors chosen and the distance norm (other than Euclidean) used for computing correspondence costs, all require further investigation. Finally, spectral methods can be sensitive to the presence of outliers in the data. However, this issue is not of great concern for 3D model retrieval, as the 3D models are mostly free of outliers. Moreover, since most models define a surface, outliers are easy to detect and remove. Study into ways of making the spectral method robust to outliers is more interesting and necessary with respect to retrieval and recognition of more general form of data.

References

1. P. Shilane, P. Min, M.K., Funkhouser, T.: The princeton shape benchmark. In: Proc. of Shape Modelling International. (2004)
2. D.-Y. Chen, X.-P. Tian, Y.T.S., Ouhyoung, M.: On visual similarity based 3d model retrieval. In: Computer Graphics Forum. (2003) 223–232
3. McGill 3D shape benchmark: http://www.cim.mcgill.ca/ shape/benchMark/.
4. Tangelder, T., Veltkamp, R.: A survey of content based 3d shape retrieval methods. In: Proc. of Shape Modeling International. (2004) 145–156
5. M. Kazhdan, T.F., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation os 3d shape descriptors. In: Symposium on Geometry Processing. (2003)
6. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the nyström method. IEEE Trans. on PAMI **26**(2) (2004) 214–225
7. Osada, R., Funkhouseri, T., Chazelle, B., Dobkin, D.: Matching 3d shapes with shape distribution. In: Proc. of Shape Modeling International. (2001) 154–166
8. Vranic, D.: An improvement of rotation invariant 3d shape descriptor based on functions on concentric spheres. In: Proc. of ICIP. (2003) 757–760
9. Shum, H.: On 3d shape similarity. In: Proc. of CVPR. (1996) 526–531
10. Kang, S., Ikeuchi, K.: Determining 3-d object pose using the complex extended gaussian image. In: Proc. of CVPR. (1991) 580–585
11. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Shape matching and anisotropy. ACM Trans. Graph. **23**(3) (2004) 623–629
12. Ohbuchi, R., Minamitani, T., Takei, T. In: Shape-similarity search of 3D models by using enhanced shape functions. (2003) 97–104
13. Hilaga, H., Shinagawa, Y., Kohmura, T., Kunii, T.K.: Topology matching for fully automatic similarity estimation of 3d shapes. In: SIGGRAPH. (2001) 203–212
14. Zhang, J., Siddiqi, K., Macrini, D., Shokoufandeh, A., Dickinson, A.: Retrieving articulated 3-d models using medial surfaces and their graph spectra. In: Int. Workshop On Energy Minimization Methods in CVPR. (2005)
15. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. In: NIPS. (2002) 857–864
16. Zhang, H., Liu, R.: Mesh segmentation via recursive and visually salient spectral cuts. In: Proc. of Vision, Modeling, and Visualization. (2005) 429–436
17. Carcassoni, M., Hancock, E.R.: Spectral correspondence for point pattern matching. Pattern Recognition **36** (2003) 193–204
18. Shapiro, L.S., M., B.J.: Feature based correspondence: an eigenvector approach. Image and Vision Computing **10**(5) (1992) 283–288
19. Jain, V., Zhang, H.: Robust 3d shape correspondence in the spectral domain. In: Proc. of Shape Modeling International. (2006) to appear
20. Elad, A., Kimmel, R.: On bending invariant signature of surfaces. IEEE Trans. on PAMI **25**(10) (2003) 1285–1295
21. Umeyama, S.: An eigen decomposition approach to weighted graph matching problem. IEEE Trans. on PAMI **10** (1988) 695–703
22. Scott, G., Longuet-Higgins, H.: An algorithm for associating the features of two patterns. Royal Soc. London **B244** (1991)
23. C. Gotsman, X.G., Sheffer, A.: Fundamentals of spherical parameterization for 3d meshes. In: ACM Transactions on Graphics (Proceedings of SIGGRAPH). (2003)
24. Gotsman, C., Karni, Z.: Spectral compression of mesh geometry. In: Computer Graphics (Proceedings of SIGGRAPH). (2000) 279–286
25. Yang, L.: k-edge connected neighborhood graph for geodesic distance estimation and nonlinear data projection. In: Proc. of ICPR. (2004)