THREE-DIMENSIONAL POINT CLOUD RECOGNITION VIA DISTRIBUTIONS OF GEOMETRIC DISTANCES

By

Mona Mahmoudi

and

Guillermo Sapiro

IMA Preprint Series # 2209

(May 2008)

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

Three-Dimensional Point Cloud Recognition via Distributions of Geometric Distances

Mona Mahmoudi and Guillermo Sapiro

Electrical and Computer Engineering, University of Minnesota
200 Union St. SE, Minneapolis, MN, 55455, USA

Abstract

A geometric framework for the recognition of three-dimensional objects represented by *point clouds* is introduced in this paper. The proposed approach is based on comparing distributions of intrinsic measurements on the point cloud. In particular, *intrinsic distances* are exploited as signatures for representing the point clouds. The first signature we introduce is the histogram of pairwise *diffusion distances* between all points on the shape surface. These distances represent the probability of traveling from one point to another in a fixed number of random steps, the average intrinsic distances of all possible paths of a given number of steps between the two points. This signature is augmented by the histogram of the actual pairwise *geodesic distances* in the point cloud, the distribution of the ratio between these two distances, as well as the distribution of the number of times each point lies on the shortest paths between other points. These signatures are not only geometric but also invariant to bends. We further augment these signatures by the distribution of a *curvature function* and the distribution of a *curvature weighted distance*. These histograms are compared using the χ^2 or other common distance metrics for distributions. The presentation of the framework is accompanied by theoretical and geometric justification and state-of-the-art experimental results with the standard Princeton 3D shape benchmark, ISDB, and nonrigid 3D datasets. We also present a detailed analysis of the particular relevance of each one of the different proposed histogram-based signatures. Finally, we briefly discuss a more local approach where the histograms are computed for a number of overlapping patches from the object rather than the whole shape, thereby opening the door to partial shape comparisons.

Key words: Point cloud data, 3D shape recognition, intrinsic distances, distributions.

Email addresses: mahmo022@umn.edu (Mona Mahmoudi), guille@umn.edu (Guillermo Sapiro).

1 Introduction and Key Contributions

Three-dimensional (3D) data is becoming more and more ubiquitous. 3D object retrieval is essential for tasks such as navigation, target recognition, and identification. In particular, point clouds are one of the most primitive and fundamental representations of 3D objects, obtained, e.g., from laser range scanners, and working directly with such representation is critical and challenging at the same time. See for example [9,15,19,22,27,30,31,33] and references therein for some of the recent works in this area. In this paper, we develop a framework for 3D object recognition from point cloud data. In particular, we introduce and exploit signatures which extract the intrinsic geometry of the 3D shapes represented by the point cloud.

The diffusion distance, [18], and the geodesic distance are two intrinsic (geometric) distances measured by paths constrained to travel on the point cloud surface of the shapes, and are the key components of the framework here proposed. The diffusion distance is related to the probability of traveling on the surface from one point to another in a fixed number of random steps, while the geodesic distance is the length of the shortest surface-path between two points.

Being invariant to bending of the surface makes these intrinsic distances natural and useful for recognition of non-rigid objects, see e.g., [5,14,16,23,38] for the use of the geodesic distance. While in order to obtain an explicit matching of the shapes, the matrices corresponding to pairwise distances need to be compared and matched [4,23], it has been, at least empirically, demonstrated that such computationally elaborate matchings can be often avoided in recognition tasks. In particular, in [3], the authors have shown that with high probability, shapes can be uniquely distinguished by the distribution of Euclidean (non-intrinsic) distances between pairs of points (samples on the shapes). The diffusion distance is equivalent to the Euclidean distance in an embedding space, as detailed in Section 2.1, which makes this argument about distributions applicable to diffusion distances in the embedding space as well. This argument combined with the need for a bending invariant signature, provides a solid reason to consider the distributions of intrinsic distances as signatures for object retrieval. Comparing the distributions of distances (or any other features), instead of applying traditional global matching methods, reduces the recognition problem to a one-dimensional comparison problem, considerably saving memory and computational time [14,16,20,24,28,37].

In real complex 3D scenarios, objects are often noisy and partially occluded or not completely scanned. It is therefore important to perform such 3D recognition robustly and from partial information (see also [13,26] for partial matching results). Graph-based methods in the object recognition literature, e.g., see [17] for Reeb graphs comparison and [29] for object recognition in videos, have been shown useful for partial matching based on local shape patches. Exploiting these graph-based matching techniques, combined with the intrinsic distance distributions here proposed, the introduced framework starts building in this direction of partial matching.

Motivated by these prior theoretical and computational results, in this paper we introduce and exploit distribution/histogram-based signatures for 3D shape recognition, and develop methods for global and local comparison between shapes represented by point clouds. The first signature we introduce is the distribution of the diffusion distance [7,18], which has not been used before for comparing 3D surfaces. This distance basically measures the probability of connectivity between points, considering all possible surface-constrained paths between them and not just the shortest one. The diffusion distance, which is easily computed from eigenvalue/eigenvector decompositions, is more robust than the natural geodesic distance to topological noise in the point cloud data, as well as topological errors created in the process of computing local neighborhoods due to the lack of connectivity information. The combination of both geodesic and diffusion distances also helps to better define these neighborhoods, as demonstrated in this paper. We also use as signatures the distribution of pairwise geodesic distances (the feature that has not been used before for point cloud data), and the distribution of the ratio between diffusion and geodesic distances. This ratio is a measure of the width of the shape in the parts connecting the two points being considered in the computation. We further introduce a measure of "centrality" for each point, which is the number of shortest paths between pairs of points that include the corresponding point, and use the distribution of this measure as an additional signature in this work. All the above signatures are not only intrinsic to the object, but invariant to bends as well. We also include the histogram of a curvature function and the distribution of a curvature weighted distance in our signatures in order to further improve the recognition performance. The relative contribution of each one of these histogram-based geometric signatures, which are all used here for the first time in a framework for *point cloud* shape recognition (and some like those associated with the diffusion distance for the first time for 3D recognition in general), is investigated in this work.

To compare these signatures for different shapes, both χ^2 and Jensen-Shannon divergence, [11], produce very good results. In particular, the results here reported based on the χ^2. These results are state-of-the-art for the standard datasets.

In addition to these global comparisons, and in order to develop a framework that is more geared toward finding local shape similarities, we also propose a method based on the computation of these signatures on "patches" of the point cloud data (see also [12,25,26]). In our approach, and following [25], we use random overlapping patches on the shape, with a control on the amount of overlap. In contrast to the more classical literature on patches, we explicitly consider their spatial relationship by using a graph-based approach.

The remainder of this paper is organized as follow: in Section 2, we discuss the basic concepts on the diffusion distance and the curvature classifier. Then, we describe the distribution signatures we develop based on these features, and the technique to compare these signatures in Section 3. Experimental results are presented in Section 4, and in Section 5, we discuss a graph-oriented local framework and conclude the work.

A preliminary version of this paper appeared at a workshop, [21]. Here we extend the framework by adding fundamental new signatures that improve the results, provide additional details, and present additional examples.

2 Basic Intrinsic Measures

2.1 Diffusion Distance

In [7,18] (see also [2] for related work), the authors introduced diffusion maps and diffusion distances as a method for data parametrization and dimensionality reduction. The diffusion distance is equivalent to the Euclidean distance in the embedding space corresponding to a mapping known as diffusion map. The diffusion distance between two points in the point cloud involves the average of all the paths of m steps connecting these two points (average probability of traveling between the points). This makes the diffusion distance a bending invariant function of the path length and the shape width between two points. Since this distance does not rely on just the shortest path between two points, it is more robust than the geodesic distance. As briefly mentioned before, in [3] the authors proved that the distribution of Euclidean distances is very informative of the shape. Combining the theory in [3] and the characteristics of diffusion distances, such as being the Euclidean distance in an embedding space and being bending invariant, makes the diffusion distance a good natural signature for non-rigid object recognition.

In order to compute the diffusion distance, we first create the affinity function $k(x, y)$ over all pairs of points x, y in the point cloud. These values become the elements of an $N \times N$ square matrix K, where N is the number of available points. This matrix is symmetric, positive semidefinite, and positive. If we then define $a(x,y)$ as

$$a(x,y) := \frac{k(x,y)}{v(x)}, \tag{1}$$

where $v(x) := \sum_y k(x,y)$ is the sum of the elements in each row, the matrix A, composed by the elements $a(x,y)$, can be viewed as the probability for a random walker on the point cloud to make a step from x to y. Now if we further define $\tilde{a}(x,y)$ as

$$\tilde{a}(x,y) := a(x,y)\sqrt{\frac{v(x)}{v(y)}}, \tag{2}$$

the corresponding matrix $\tilde{A}$ is symmetric and can be decomposed as

$$\tilde{a}(x,y) = \sum_{i=0}^{N} \lambda_i^2 \phi_i(x)\phi_i(y), \tag{3}$$

where $\lambda_0^2 = 1 \geq \lambda_1^2 \geq \lambda_2^2 \geq ... \geq \lambda_N^2$ are the eigenvalues (note the "square," which will simplify the expressions later), of the matrix $\tilde{A}$ and ϕ_l are the corresponding eigenvectors. Therefore, for the elements of the matrix $\tilde{A}^m$ we obtain

$$\tilde{a}^{(m)}(x,y) = \sum_{i=0}^{N} \lambda_i^{2m} \phi_i(x)\phi_i(y), \tag{4}$$

which can be interpreted as representing the probability for a random walker or Markov chain with transition matrix $\tilde{A}$ to reach y from x in m steps.

Following in part standard concepts from kernel methods, the authors in [7] introduced the diffusion map (Φ_m) from the given point cloud data to an Euclidean space using the kernel $\tilde{a}^{(m)}$. This mapping is obtained as

$$\Phi_m(x) = \begin{pmatrix} \lambda_0^m \phi_0(x) \\ \lambda_1^m \phi_1(x) \\ \lambda_2^m \phi_2(x) \\ \vdots \end{pmatrix}. \tag{5}$$

It is easy to prove, e.g., see [34] for more details on these kernel methods, that the Euclidean distance between the mapped points $\Phi_m(x)$ and $\Phi_m(y)$ in the new space is

$$D_m^2(x,y) = \tilde{a}^{(m)}(x,x) + \tilde{a}^{(m)}(y,y) - 2\tilde{a}^{(m)}(x,y), \tag{6}$$

which is exactly the diffusion distance between points x and y. (The selected values for m and other parameters are presented in Section 4.)

In order to separate the geometry of the point cloud from its density, $k(x,y)$ is further normalized, [18],

$$\tilde{k}(x,y) := \frac{k(x,y)}{p(x)p(y)}, \tag{7}$$

where $p(x) := \sum_y k(x,y)$, and $\tilde{k}$ is used in Eq. (1) instead of k.

In this work, we first use the Gaussian kernel $k(x,y) = \exp(-||x-y||^2/\sigma^2)$ to define the affinity matrix, where σ is the average of Euclidean distances between all pairs of points in the shape. As a result of using Euclidean distances to define this affinity kernel, we have topological shortcuts in computing the diffusion distance. This is illustrated in Figure 1, where some points on the legs of the dog are so close to each other in terms of Euclidean distance that the "shortcut" leads to an undesired (and incorrect) small diffusion distances between the two adjacent back legs. One possible solution would be to reduce the value of σ. However, with a small σ, many points become isolated and their diffusion distance to all other points becomes too large. To avoid such shortcuts, we first compute the geodesic distance between all the points in the shape, computation done using Floyd's algorithm on the graph obtained from connecting only a few nearest neighbors, 3-6 neighbors in our case (an alternative technique is given in [22] which works directly on the point cloud). Then, for each point x we find the set $\mathcal{M}(x)$ of g-nearest neighbors of x, in terms of geodesic distance. Then, we define $k(x,y)$ by a neighborhood filtering as

$$k(x,y) = \begin{cases} e^{(-\frac{||x-y||^2}{\sigma^2})} & y \in \mathcal{M}(x), \\ 0 & y \notin \mathcal{M}(x). \end{cases} \tag{8}$$

See in Figure 1 how this addresses the shortcuts problem.

This concludes the presentation of the diffusion distance, and we now proceed to present the basic concepts of the curvature classifier.

Fig. 1. *In both pictures the colors show the diffusion distance for all the points from a fixed point in one of the legs of the dog (dark blue for small and dark red for large values). The left picture shows the case without neighborhood filtering in computing the diffusion distance, obtaining undesired shortcuts (see how the back legs are considered close). In the right figure we observe how these shortcuts are avoided by using the neighborhood filtering based on the geodesic distance. (This is a color figure.)*

2.2 Curvature Classifier

We now describe a local surface classifier introduced in [6], which will be used to augment the discriminatory power of the diffusion and geodesic distances. This classifier robustly distinguishes between smooth regions and edges or corners. While the distributions of intrinsic distances and their ratio ignore small parts on the shape which have high curvature, using the distributions of a function of the curvature and a curvature weighted distance, as additional signatures, helps in recognizing these parts.

If M is the considered surface and $B_\epsilon(x)$ is an Euclidean ball with radius ϵ centered at a point x, we define the zero moment of the ϵ-neighborhood of x as

$$M_\epsilon^0(x) := \int_{B_\epsilon(x)\cap M} x dx, \tag{9}$$

and its first moment as

$$\begin{aligned} M_\epsilon^1(x) &:= \int_{B_\epsilon(x)\cap M}(x - M_\epsilon^0(x)) \otimes (x - M_\epsilon^0(x))dx \\ &= \int_{B_\epsilon(x)\cap M} x \otimes x - M_\epsilon^0(x) \otimes M_\epsilon^0(x)dx, \end{aligned} \tag{10}$$

where $y \otimes z := (y_i z_j)_{i,j=1,2,3}$. These moments are expected to be robust to noise, and provide information about the curvature at x, using the eigenvalues of the first moment and the zero moment shift defined as

$$T_\epsilon(x) := M_\epsilon^0(x) - x. \tag{11}$$

For example, $T_\epsilon(x)$ scales quadratically with the filter width ϵ in smooth areas and linearly at corners and edges. The following function of these moments is then used as a measure of curvature:

$$C_\epsilon := G\left(\frac{||T_\epsilon||\lambda_{min}}{\epsilon\lambda_{max}}\right), \tag{12}$$

where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of the first moment at point x, respectively. In particular, we consider $G(s) = \frac{1}{\alpha+\beta s^2}$, with appropriately chosen α and β. In our application, we have set $\alpha = .002$ and $\beta = 2000$. The value of $C_\epsilon(x)$ will be close to $\frac{1}{\alpha}$ at smooth areas and $C_\epsilon(x) << \frac{1}{\alpha}$ at corners and edges.

Having the basic concepts of intrinsic distances and curvature functions, we now proceed to present the signatures derived from them and the proposed recognition framework.

3 Recognition Framework

In this section, we present the signatures we use in order to recognize 3D objects represented by point clouds, and the techniques for comparing between these signatures in different shapes.

3.1 Characterizing Signatures

In this part, we present six characterizing signatures which, except for the histogram of the geodesic distance, which has been used but for meshes, have not been previously used in 3D object recognition.

Histogram of diffusion distance. As our first signature, we use the histogram of diffusion distance, motivated by the discussion in Section 2.1. Being bending invariant, similar to geodesic distance, it has the advantage of being more robust to noise since it exploits all the paths of fixed number of steps, not only the shortest one as in geodesic distance.

Histogram of geodesic distance. As mentioned above, the geodesic distance is the length of the shortest path, constrained to the manifold, between two points. Works such as those in [14,16] have used the histogram of the average geodesic distance from a point to the rest as a signature for shape recognition (primarily for meshes). This is motivated in part by the fact that geodesic distances are the basic bending invariant features of the shape, and thereby useful for non-rigid object recognition [23]. When compared with the diffusion distance, the geodesic distance is more sensitive to noise, and it is thereby used here to augment the other features, and not alone. We compute this distance by Floyd's algorithm, while we could also use the work in [22] to compute it directly on the point cloud. To avoid shortcuts, we start with three nearest neighbors in the neighborhood graph and increase it by one in each step, until the constructed graph is connected or it reaches a maximum number.

Histogram of the ratio between diffusion and geodesic distances. The diffusion distance contains information about the "width" of the object in the area connecting two points by considering the number of paths with a fixed number of steps between them, in addition to their distance on the manifold. Since the geodesic distance is the length of the shortest path between two points, the ratio between the diffusion distance and the geodesic distance provides information about the average width in the path between the two points. The histogram of this ratio is the third signature considered here. Since for small geodesic distances, the ratio is too large, we have excluded the distances that are smaller than a threshold. The threshold we use in our experiments is three times the average of the smallest nonzero geodesic distance at each point, and we remove all pairs of points with a geodesic distance less than this threshold.

Histogram of a centrality measure. One of the characteristics of a point in a 3D surface is its intrinsic centrality. We propose a new function to measure the centrality of each point, which is the number of shortest paths (geodesic curves) between all the pairs of points in the shape that include the specific point (to avoid noise and the possible effects of non-uniformity of the samples, we can average this number in a K-neighborhood of each point). We expect higher values of this measure for points closer to the center or in the center of narrow parts (for example, legs of the animals), and lower values for the end points. In the proposed point cloud recognition framework, the histogram of this measure for all the points in a shape is used as an additional signature.

Histogram of the curvature classifier. In our experiments, we noticed that considering only the histograms of bending invariant distances neglects the information in the small high curvature parts. This becomes more critical for recognizing classes of 3D objects as in the results presented in Section 4, and not just single bended representatives per class as in [10,23]. For this purpose, we propose two additional new signatures, the histogram of the curvature classifier described in Section 2.2, and the histogram of a curvature weighted distance (see below for the description of this signature). Since there are a lot of low curvature points in each shape and many high curvature parts are caused by noisy or non-smoothly sampled manifolds, the part of the curvature histograms corresponding to these very low or high curvature points is not informative. Thus, disregarding them improves the results.

Histogram of a curvature weighted distance. Following the above discussion about considering curvature as a distinguishing feature, we define a new distance between points which gives larger weights to the points with higher curvature. This curvature weighted distance is computed by accumulating a linear decreasing function of the curvature classifier, explained in Section 2.2, over the shortest paths between all pairs of points (natural geodesic). We use the histogram of these distances as the last of the proposed signatures.

In Figure 2 we illustrate the diffusion distance, geodesic distance, their ratio, the curvature weighted distance, and the curvature classifier, as well as the centrality measure, for a few examples. In Figure 3, we present each one of the six distribution/histogram-based signatures for some representative shapes.

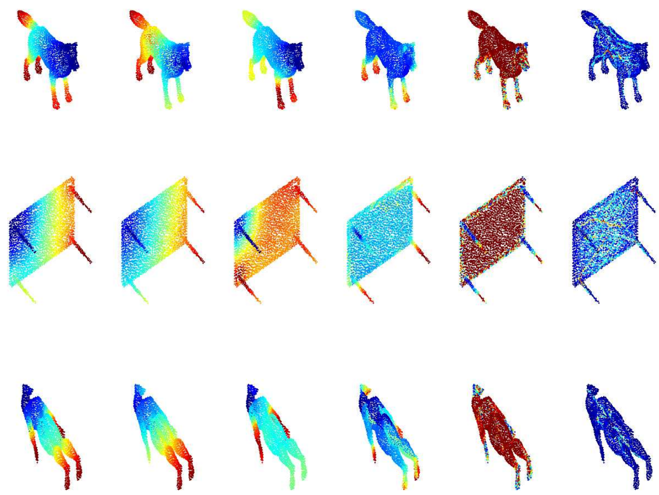

Fig. 2. *From left to right in each row: The value of the diffusion distances, geodesic distances, their ratio, and the curvature weighted distance, from a point (dark blue) to the rest of the 3D shape; followed by the value of the curvature classifier and the centrality measure for all points. Dark blue represents small values and dark red large values. (This is a color figure.)*

3.2 Signatures Comparison

To conclude the description of the global shape recognition framework, we must describe how we combine and compare the above mentioned histograms. In order to compare two histograms, which are automatically normalized to compensate for the shape scale, we tested different distance measures, such as L_1 and L_2 norms, χ^2, correlation coefficients, and the Jensen-Shannon diver-

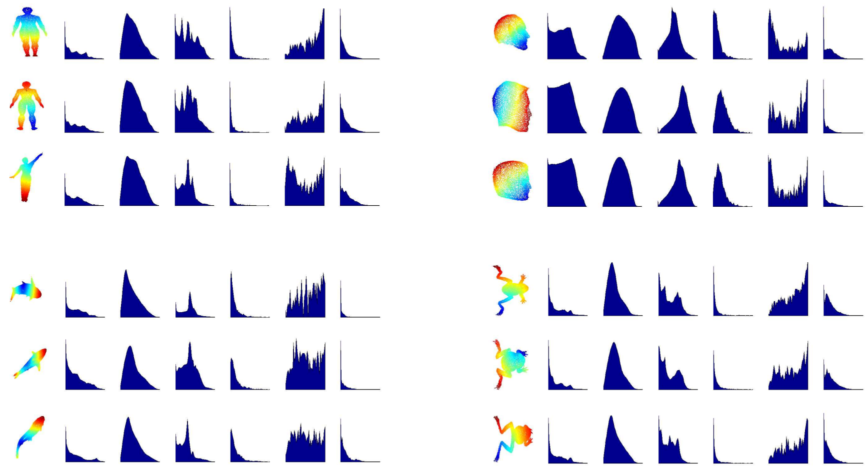

Fig. 3. *All six histograms are shown following the respective shapes. The histograms, from left to right, are presented in the order described in the text. Colors on the shapes correspond to the geodesic distance from one point on the shape to the rest. (This is a color figure.)*

gence (JSD), which is the symmetric and smoothed version of the Kullback-Leibler divergence. The best results were obtained for the χ^2 measure, followed by the JSD. Therefore, in our results, we have used the χ^2 measure between two normalized Z-bins distributions, h_i and h_j, which is given by

$$\delta_{ij} := \frac{1}{2}\sum_{z=1}^{Z}\frac{(h_i(z)-h_j(z))^2}{h_i(z)+h_j(z)}.$$

Having the basic way to compare pairs of histograms, now we need to combine the distance metric for the six signatures presented in the previous section in order to obtain the "dissimilarity" between two shapes. For the results presented in Section 4, we multiply the six distances obtained for each one of the six different histograms. This leads to better results than, for example, considering multidimensional histograms of two or more features.

As detailed in the next section, this simple distance between histograms already leads to state-of-the-art results. In the future we plan to further investigate replacing the χ^2 by other metrics, and also other ways of combining the signatures, including automatically learning the weights and relevance of each one of them.

4 Experimental Results

In our experimental results, for comparison, we use 3D shapes from the same database tested in [14], which is the combination of two different databases, part of the Princeton Shape Benchmark (PSB), [35], and ISDB. These two databases consist of 22 categories, overall. We also present the results tested on a nonrigid 3D database (NR) [4].

We have a total of 635 shapes from 27 categories in the three databases. Since our proposed recognition techniques do not rely on the connectivity information in these triangulated data, we first converted them to point clouds. We have uniformly sampled 3000 random points from vertices of each shape, using the maxmin sampling method in [8], after subdividing the triangles using the Graph toolbox in MATLAB [32]. Even if the point samples of a shape are non-uniform but large enough, 3000 points can be uniformly sampled without loss of generality for originally non-uniform point clouds. We have used $m = 50$ for the number of steps of the path in the diffusion distance, $g = 100$ nearest neighbors to find $\mathcal{M}(x)$ in Eq. (8), and only the 6 largest eigenvalues of $\tilde{A}$. In computing the curvature function, we used 8 nearest neighbors for each point and defined ϵ as the maximum Euclidean distance to the 8-th neighbor of all points. Since the maximum value of the curvature classifier is 500 (based on the selected values of α and β), in computing the curvature weighted distance, we use the curvature classifier subtracted from 500 at each point, as the actual curvature function. All six histograms have 50 bins. For the curvature classifier histogram, considering only the last 40 bins leads to better recognition, as discussed in Section 3.1.

In order to evaluate the effectiveness of the different signatures and methods, we first find the similarity measure between each pair of shapes by applying each signature to form a square matrix of dissimilarity values. We use the following three criteria for the recognition performance:

Nearest neighbor: The percentage of the cases where the query belongs to the same category as its closest match (not considering the query itself).

First tier: The percentage of the shapes in the same category as the query that are among its U closest matches, where $U + 1$ is the total number of shapes in the corresponding category.

Second tier: This value is the same as in the first tier with the difference that now the $2U$ closest matches are considered.

The percentages presented here are the average values of these measures over one category or the whole dataset. Although, the commonly used first and second tiers are good criteria for evaluating recognition methods when the intraclass variability is low, it can be a misleading measure when there is a lot of variability in the classes. For example, based on the signatures used for recognition, a square chair without handle can be more similar than a round table to a square table. In this case, assume we have the same number of square chairs, square tables, and round tables in the database; and the tables are all in the same category. In this situation, since the chairs are closer matches to square tables than the round tables are, the first tier of square tables can be close to 50%. On the other hand, if the tables are categorized in two different categories, with exactly the same signatures, the first tier for square tables can be increased to 100%. This discussion shows that the amount of intraclass variability in different databases makes a big difference in values of tiers. The PSB database, that is reported here, has a large intraclass variability in many of the categories, and the ISDB and NR databases have lower variability within each class. Thus, lower values for tiers is expected in the PSB dataset.

In Table 1, the results of using each signature as well as some combinations of them over the three datasets are presented. For comparison, we also included the results obtained when using the histogram of the average geodesic distance from each point to every other point, which was used in [14,16] for meshes. Among the single signature methods, the best result, considering the best match, is obtained by the proposed diffusion distance, and the best overall result is obtained by combining our proposed six signatures. In Figure 4, the best matches given by combining these six signatures are presented for six representative shapes.

In Table 2, the results for some of the objects categories by using the proposed global comparison (DCRGcDP) and the state-of-the-art CDF method [14], over the whole dataset used in [14], are presented. In the table, we present the results for some of the 22 classes, containing all the ones reported in [14]. In [14], the authors use, as the signature, a two dimensional histogram of the combination of the average geodesic distance (which as shown in Table 1 is not as good as diffusion distance), and a measure of diameter of the shape around each point over the triangulated data.

We observe that the overall performance of our method, considering the best match, over the whole dataset is better than the performance of the CDF method, which reported state-of-the-art results at the time of publication. One can observe that in both techniques, the categories of “humans,” “horses,” “human hands,” and “furniture” have the highest correct recognition rates. We have noticeably better results in categories of “airplanes,” “humans,” “ships,” “furniture,” and “fishes,” showing that our proposed descriptors better capture the intrinsic characteristics of those classes. Having a very diverse collection of models, the classes “chairs,” “tables,” “insects,” and “helicopters” show lower performance. Finally, note that unlike most algorithms reported in the literature, including [14], we do not rely on the neighborhood information in

	Total (635)			ISDB (106)			PSB (381)			NR (148)		
	BM	FT	ST	BM	FT	ST	BM	FT	ST	BM	FT	ST
D	68%	32%	47%	92%	59%	68%	56%	29%	44%	89%	63%	81%
G	57%	32%	48%	74%	45%	65%	52%	32%	47%	79%	45%	66%
mG	52%	29%	45%	82%	52%	73%	47%	29%	44%	73%	42%	65%
R	57%	26%	42%	75%	45%	60%	43%	24%	38%	86%	58%	74%
C	43%	28%	43%	69%	49%	68%	40%	23%	37%	81%	57%	76%
cD	50%	24%	38%	71%	45%	63%	35%	20%	32%	84%	55%	72%
P	28%	19%	33%	52%	36%	56%	31%	23%	39%	24%	20%	41%
DC	72%	35%	51%	91%	65%	74%	60%	30%	46%	91%	67%	83%
DR	66%	32%	46%	87%	56%	67%	51%	29%	43%	88%	63%	80%
DG	71%	39%	55%	89%	59%	71%	62%	35%	50%	91%	66%	82%
DRC	73%	35%	51%	92%	63%	73%	60%	32%	47%	90%	66%	82%
DGR	74%	38%	54%	88%	60%	71%	62%	36%	49%	94%	66%	82%
DCRG	78%	40%	56%	91%	65%	75%	68%	38%	51%	95%	69%	83%
DCRGcD	79%	38%	54%	94%	68%	77%	68%	36%	50%	95%	69%	82%
DCRmGcDP	78%	35%	50%	97%	71%	79%	68%	33%	49%	93%	67%	82%
DCRGcDP	80%	38%	53%	95%	70%	79%	71%	37%	51%	95%	69%	82%

Table 1
Effectiveness of each one of the six signatures (plus average geodesic) and some of their combinations, evaluated with the global comparison method over the three datasets: ISDB, PSB, NR, and the combination of all of them (Total). In the table, D stands for diffusion distance, G for geodesic distance, mG for average geodesic distance, R for ratio of diffusion and geodesic distances, P for the centrality signature, C for curvature classifier, and cD for the curvature weighted distance. The evaluation measures presented here are best match (BM), first tier (FT), and second tier (ST).

the triangulated data. This lack of information leads to lower recognition in some categories, for example "cats," when the cat is seated. Overall, recognizing point-clouds is significantly more challenging than working with meshes, while we still obtain state-of-the-art results when compared to mesh-based approaches.

In Table 3, values of best match, first tier, and second tier are presented for the databases ISDB and PSB and their combination (Total), for four methods: DCRGcDP, CDF, Light Field Descriptor (LFD), and Spherical Harmonics (SH), based on the results reported in [14]. Light Field Descriptor (LFD) and Spherical Harmonics (SH) are two out of the three top performing descriptors for PSB as described in [35]. Among all the four methods, we have the second best overall results, considering the best match. As discussed above, the tiers corresponding to PSB database with larger amount of intraclass variability are lower than the tiers corresponding to the ISDB database. Recall that our results are for the more challenging point cloud 3D data representation, while the other algorithms are reported on meshes.

		Best Match	First Tier	Second Tier
Planes (29)	DCRGcDP	76%	26%	37%
	CDF	45%	25%	44%
Humans (134)	DCRGcDP	99%	60%	89%
	CDF	88%	57%	84%
Horses (16)	DCRGcDP	88%	64%	73%
	CDF	94%	68%	85%
Hands (33)	DCRGcDP	85%	48%	58%
	CDF	82%	67%	76%
Insects (20)	DCRGcDP	60%	17%	24%
	CDF	71%	23%	34%
Chairs (33)	DCRGcDP	58%	18%	30%
	CDF	45%	20%	34%
Ships (21)	DCRGcDP	67%	19%	25%
	CDF	24%	11%	20%
Guns (7)	DCRGcDP	71%	29%	35%
	CDF	71%	40%	50%
Furniture (19)	DCRGcDP	89%	37%	56%
	CDF	63%	NA	NA
Fishes (26)	DCRGcDP	81%	37%	48%
	CDF	65%	NA	NA
Birds (20)	DCRGcDP	40%	18%	23%
	CDF	30%	NA	NA
Total (487)	DCRGcDP	76%	38%	53%
	CDF	71%	45%	63%

Table 2
Recognition results for both DCRGcDP and CDF matching methods for some of the 3D object categories, where the recognition is among all the shapes in PSB and ISDB datasets used in [14].

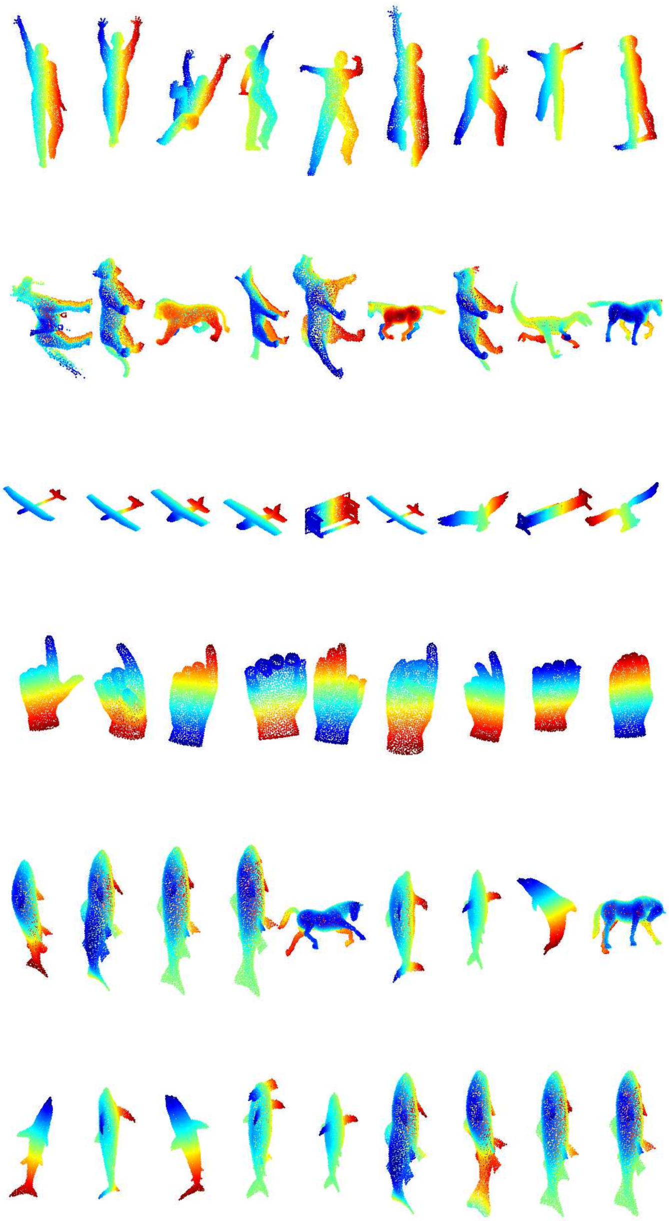

Fig. 4. *Results of shape retrieval for the global recognition algorithm using all six histogram-based signatures. The first column on the left shows the query models, and the other figures on each row show the top eight matches. (This is a color figure.)*

5 Discussions, Local Analysis, and Conclusions

In this paper, we introduced a new framework for 3D object recognition from point cloud data. The proposed 3D signatures are derived from the distribution of the pairwise diffusion distances, the distribution of the pairwise geodesic distances, the distribution of the ratio between these two distances, the distribution of a centrality measure, the distribution of a curvature classifier, and the distribution of a curvature

	Total (487)			ISDB (106)			PSB (381)		
	BM	FT	ST	BM	FT	ST	BM	FT	ST
DCRGcDP	76%	38%	53%	95%	70%	79%	71%	37%	51%
CDF	71%	45%	63%	100%	98%	100%	65%	40%	58%
LFD	79%	42%	59%	73%	44%	62%	87%	47%	62%
SH	75%	37%	54%	78%	47%	64%	77%	41%	57%

Table 3
The overall results of the four methods, DCRGcDP, LFD, SH, and CDF, tested on ISDB and PSB databases and their combination (Total) is presented.

weighted distance. The use of intrinsic distances and their distributions is supported by theoretical work as well as by extensive experimental results in both the 3D shape recognition and image analysis literature. Although the distribution of geodesic distances has been used before for 3D recognition of triangulated surfaces (not point clouds as here reported), the other signatures have not been incorporated in prior art.

Since the information in the signatures (histograms) defined on the whole shape is global, it might ignore some important local information for identification. It is thereby reasonable to compute the signatures more locally. In addition, in practical scenarios where occlusions (or partial acquisition) are present, there is a need for more local signatures. We extend the global framework to (semi-)local recognition by considering overlapping patches (similar to the idea in [25]). Patches, originally, are 50 sets of the 300 closest, in the geodesic sense, points to 50 center points, sampled from the shape by the maxmin sampling method [8]. Then, all the patches with more than 70% overlap are joined as one patch. These patches become nodes in a graph, with attributes given by the six histograms described in Section 3.1, and edges encoding the spatial relationship between the patches (connecting the nodes corresponding to two neighboring patches). The edge weights are the geodesic distances between the two corresponding center points, computed on the whole shape. Then, we apply a graph comparison algorithm, following in part the work introduced in [29] for shape recognition in video. We have applied this method over a dataset of 119 shapes from the Princeton Shape Benchmark (PSB), [35], and SCAPE pose and body shapes data [1], and the preliminary overall obtained results where comparable to the global point cloud 3D shape recognition method introduced in this paper. In categories such as tables, human hands, and insects, the graph method produced better results; while for cars, planes, and horses, the global method lead to better results. One of our ongoing objectives is to further improve the graph comparison method and to use it in partial matching applications.

We are also considering combining the framework here proposed with topological techniques, e.g., [36], in particular to address diverse classes such as chairs.

We have started to experiment with more advanced classification methods from the learning community, applying them to our signatures, e.g., SVM, which have been very successfully used in the image recognition literature. Preliminary results are encouraging, since straightforward use of SVM produces similar results to the χ^2 metric. Results in all these direction will be reported elsewhere.

Acknowledgments

We thank R. Gal for some clarifications about his work and the people behind the different datasets used in this paper for providing them. This work was partially supported by ARO, NSF, ONR, NGA, and DARPA.

References

[1] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis, "SCAPE: shape completion and animation of people," *ACM Trans. Graphics* **24-3**, pp. 408-416, July 2005.

[2] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation* **15-6**, pp. 1373-1396, June 2003.

[3] M. Boutin and G. Kemper, "On reconstructing n-point configurations from the distribution of distances or areas," *Adv. Appl. Math.* **32-4**, pp. 709-735, May 2004.

[4] A. Bronstein, M. Bronstein, and R. Kimmel, "Efficient computation of isometry-invariant distances between surfaces," *SIAM Journal of Scientific Computing* **28-5**, pp. 1812-1836, September 2006.

[5] A. Bronstein, M. ronstein, and R. Kimmel, "Robust expression-invariant face recognition from partially missing data," *Proc. European Conf. Computer Vision (ECCV)*, Graz, Austria, May 7-13, 2006.

[6] U. Clarenz, M. Griebel, M. Rumpf, A. Schweitzer, and A. Telea, "Feature sensitive multiscale editing on surfaces," *Visual Computer* **20-5**, pp. 329-343, July 2004.

[7] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, **21**, pp. 5-30, July 2006.

[8] V. de Silva and G. Carlsson, "Topological estimation using witness complexes," *Proc. Sympos. Point-Based Graphics*, 2004.

[9] N. Dyn, M. S. Floater, and A. Iske, "Adaptive thinning for bivariate scattered data," *Journal of Computational and Applied Mathematics* **145-2**, pp. 505-517, August 2002.

[10] A. Elad (Elbaz) and R. Kimmel, "Bending invariant representations for surfaces," *Proc. Computer Vision and Pattern Recognition (CVPR)*, pp. 168-174, December 2001.

[11] B. Fuglede and F. Topse, "Jensen-Shannon divergence and Hilbert space embedding," *Proc. IEEE Int'l Symp. Information Theory*, June 27-July 2, 2004.

[12] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, "A search engine for 3D models," *ACM Trans. Graphics* **22-1**, pp. 83-105, January 2003.

[13] T. Funkhouser and P. Shilane, "Partial matching of 3D shapes with priority-driven search," *Eurographics Symp. Geometry Processing*, June 26-28, 2006.

[14] R. Gal, A. Shamir, and D. Cohen-Or, "Pose-Oblivious Shape Signature," *IEEE Trans. Visualization and Computer Graphics* **13-2**, pp. 261-271, 2007.

[15] M. Gross, H. Pfister, M. Alexa, M. Pauly, and M. Stamminger, "Point based computer graphics," *EUROGRAPHICS Lecture Notes* , 2002.

[16] A. B. Hamza and H. Krim, "Probabilistic shape descriptor for triangulated surfaces," *Proc. IEEE Int'l Conf. Image Processing (ICIP)* **1**, pp. 1041-1044, September 11-14, 2005.

[17] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii, "Topology matching for fully automatic similarity estimation of 3D shapes," *Proc. SIGGRAPH 2001, Computer Graphics Proc., Annual Conf. Series*, pp. 203 - 212, August 2001.

[18] S. Lafon, "Diffusion maps and geometric harmonics," *Ph.D. dissertation*, Yale University, 2004.

[19] L. Linsen and H. Prautzsch, "Local versus global triangulations," *EUROGRAPHICS '01*, 2001.

[20] D. G. Lowe, "Distinctive image features from scale invariant keypoints," *Int'l Journal Computer Vision* **60-2**, pp. 91-110, November 2004.

[21] M. Mahmoudi and G. Sapiro, "Three-dimensional point cloud recognition via distributions of geometric distances," *IEEE CVPR Workshop on Search in 3D*, Alaska, June 2008.

[22] F. Memoli and G. Sapiro, "Distance functions and geodesics on submanifolds of R^d and point clouds," *SIAM Journal Applied Math.* **65-4**, pp. 1227-1260, October 2005.

[23] F. Memoli and G. Sapiro, "A theoretical and computational framework for isometry invariant recognition of point cloud data," *Foundations of Computational Mathematics* **5-3**, pp. 313-347, July 2005.

[24] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence* **27-10**, pp. 1615-1630, October 2005.

[25] N. J. Mitra, L. Guibas, J. Giesen, and M. Pauly, "Probabilistic Fingerprints for Shapes," *Eurographics Symp. Geometry Processing*, June 26-28, 2006.

[26] N. J. Mitra, L. Guibas, and M. Pauly, "Partial and approximate symmetry detection for 3D Geometry," *ACM Trans. Graphics* **25-3**, pp. 560-568, July 2006.

[27] N. J. Mitra, A. Nguyen, and L. Guibas, "Estimating surface normals in noisy point cloud data," *Int'l Journal of Computational Geometry and Appl.* **14-(4-5)**, pp. 261-276, October 2004.

[28] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Trans. Graphics (TOG)*, 2002.

[29] K. A. Patwardhan, G. Sapiro, and V. Morellas, "A graph-based foreground representation and its application in example based people matching in video," *IEEE Int'l Conf. Image Processing*, September 2007.

[30] M. Pauly and M. Gross, "Spectral processing of point-sampled geometry," *Computer Graphics (SIGGRAPH 2001 Proc.)*, pp. 379-386, August 12-17, 2001.

[31] M. Pauly, M. Gross, and L. Kobbelt, "Efficient simplification of point-sampled surfaces," *Proc. IEEE Visualization*, Boston, USA, pp. 163-170, October 27-November 1, 2002.

[32] G. Peyré, "Graph theory toolbox," `www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=5355\&objectType=file`, 2007.

[33] S. Rusinkiewicz and M. Levoy, "QSplat: A multiresolution point rendering system for large meshes," *Computer Graphics (SIGGRAPH 2000 Proc.)*, pp. 343-352, July 23-28, 2000.

[34] B. Schölkopf and A. Smola, *Learning with Kernels. Support Vector Machines, Regularization, Optimization and Beyond,* The MIT Press, Cambridge, 2002.

[35] P. Shilane, M. Kazhdan, P. Min, and T. Funkhouser, "The Princeton shape benchmark," *Proc. Int'l Conf. Shape Modeling and Applications(SMI)*, Genoa, Italy, pp. 167-178, June 2004.

[36] G. Singh, F. Memoli, and G. Carlsson, "Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition," *Point Based Graphics*, Prague, Czech Republic, September 2007.

[37] J. Sivic and A. Zisserman, "Video Google: Efficient visual search of videos," *Toward Category-Level Object Recognition, Springer*, 2006.

[38] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science* **290-5500**, pp. 2319-2323, December 2000.