3D Model Retrieval With Morphing-Based Geometric and Topological Feature Maps

Meng Yu, Indriyati Atmosukarto, Wee Kheng Leow, Zhiyong Huang, Rong Xu
Dept. of Computer Science, National University of Singapore
3 Science Drive 2, Singapore 117543
yumeng, indri, leowwk, huangzy, xurong@comp.nus.edu.sg

Abstract

Recent advancement in 3D digitization techniques have prompted to the need for 3D object retrieval. Our method of comparing 3D objects for retrieval is based on 3D morphing. It computes, for each 3D object, two spatial feature maps that describe the geometry and topology of the surface patches on the object, while preserving the spatial information of the patches in the maps. The feature maps capture the amount of effort required to morph a 3D object into a canonical sphere, without performing explicit 3D morphing. Fourier transforms of the feature maps are used for object comparison so as to achieve invariant retrieval under arbitrary rotation, reflection, and non-uniform scaling of the objects. Experimental results show that our method of retrieving 3D models is very accurate, achieving a precision of above 0.86 even at a recall rate of 1.0.

1. Introduction

Recent developments in 3D shape modeling and digitization techniques have led to an increased accumulation of 3D models in databases and on the Internet, and has prompted to the need of developing effective techniques for retrieving similar 3D objects given a query model. The fundamental problem of 3D object retrieval, similar to 2D image retrieval, is the measurement of similarity between two objects. A well defined similarity measure should reflect the nature of the similarity between two objects qualitatively and quantitatively.

Our shape comparison method is based on 3D morphing. Given two 3D objects, morphing involves the production of a sequence of intermediate objects that gradually change from one object to the other [3]. The *amount of effort* required to morph an object into another object can be used to measure the difference between them. Instead of morphing a source object to a target object, we can also morph both objects into a canonical object, such as a sphere or a cylinder, and compare the difference between the 3D objects and the canonical object.

3D objects can be characterized by two main features: geometry and topology. Intuitively, geometry determines the shapes of the objects and shape features such as size, curvature, and smoothness of object surfaces. On the other hand, topology determines the structures of the objects such as the number of holes and disconnected components.

To capture both geometric and topological features, two spatial feature maps are used in our method based on the idea of 3D morphing, without performing explicit morphing operation. Given a 3D object, it is first scaled and embedded in a sphere of unit radius such that the center of the sphere coincides with the object's centroid. Then, a ray is shot from the center of the sphere through each point of the object to the sphere's surface (Fig. 1). The distance traveled by the ray from an object point to the sphere's surface is recorded in the *Distance Map* (DM).

Note that, being a spatial map, DM captures not only the distances corresponding to the points, but also the spatial locations of the points. So, it is different from the distance distributions or histograms used in existing methods (e.g., [21, 22]), which do not capture spatial information of the object points. In essence, DM measures the displacements, which are proportional to the energy, required to morph an object into a sphere. It is a measure of the geometrical difference between the object and the sphere.

The number of surfaces penetrated by the ray as it travels from the sphere's center through each object point to the sphere's surface is recorded in the *Surface Penetration Map* (SM). It measures the number of surfaces that must be compressed into a single surface to morph the object into a sphere. It describes the topology and concavity of the object in terms of the number of surfaces that result from holes, disconnected components, as well as concave parts of the object. Details about the extraction of these geometric and topological feature maps are discussed in Section 3 and experimental tests and results are described in Section 4.

The following characteristics distinguish our method from existing ones:

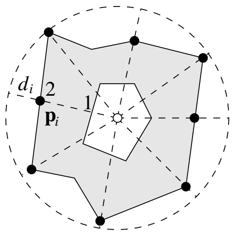

Figure 1. Computing feature maps. Rays (dashed lines) are shot from the center (white dot) of a bounding sphere (dashed circle) through the object points (black dots) to the sphere's surface. The distance d_i traveled by the ray from a point $\mathbf{p}_i$ to the sphere's surface and the number of object surfaces (solid lines; 2, in this case) penetrated by the ray since it leaves the sphere's center are recorded in the feature maps.

- Our method is based on 3D morphing, and the amount of effort required to morph 3D objects into a canonical object is used to measure their similarity.
- It uses feature maps that capture spatial information about the features. In contrast, many existing methods (e.g., [1, 2, 9, 11, 12, 13, 14, 16]) use histograms of shape features that do not capture spatial information.
- It explores the use of both geometric features and topological features. On the other hand, almost all existing methods use geometric features only, except [6] which uses topological feature.

2. Related Work

The use of canonical object for shape comparison has been applied by Hebert et al. [5] to object recognition. Their method deforms the mesh representation of an ellipsoid onto a 3D object and measures the simplex angle at each node of the mesh. The difference between two 3D objects is computed by comparing the nodes angles in their meshes. This method is applicable only to 3D objects that are topologically equivalent and geometrically similar to a sphere.

The method of Hilaga et al. [6] is the only method that uses topological feature for 3D object matching. The topology of an object is represented in a *reeb graph*. The computation of the reeb graph requires vertex resampling, short-cut edge generation, and computation of the geodesic distance. In contrast, the topological feature used in our method is simpler and far less expensive to compute than the reeb graph, as will be evident in Section 3.3.

Existing methods that use geometric features for 3D object retrieval can be divided into three broad categories according to the type of shape features used: (1) global features, (2) histograms, and (3) spatial maps. Global features refer to shape features such as moments, aspect ratio, volume-to-surface ratio, etc. Since single feature values are used to characterize the overall shape of the objects, these features tend to be not very discriminative about the objects. They have been used in [4, 21, 22].

Histograms of local shape features are probably the most widely used feature types for 3D object retrieval. The term "histogram" has been used by various authors to mean somewhat different things. Here, we use the term to mean a discrete frequency or probability distribution of features such that each bin of a histogram represents a range of feature values and each bin count is either a frequency or a probability of occurrence of the feature values within the range of the bin. Thus, histograms capture the distribution of features over the entire object without representing spatial information of the features. In general, histograms are invariant to rotation, reflection, and uniform scaling of objects. Histograms of various feature types have been used, such as angle, distance, area, volume, and curvature [11, 12, 13, 16, 21, 22]. Special types of histograms such as *spin image* and *shape context* have also been used to represent the relative positions of points [1, 2, 9, 14].

Spatial maps are representations that capture the spatial information of an object's features. The map entries correspond to physical locations or sections of an object, and are arranged in a manner that preserves the relative positions of the features in the object. For example, Kriegel et al. [7, 8] and Suzuki et al. [17, 18] divided an object into cells and used the number of points within each cell as the feature. Vranic et al. [19, 20] computed 2D maps of spherical harmonics coefficients and Novotni and Klein [10] computed 3D maps of distances to features on the objects.

Since spatial maps preserve the spatial information of the features in an object, they are generally not invariant to linear transformations, except for specially designed maps (e.g., the rotationally invariant map of [7, 8]). So, Fourier transform is often performed to transform spatial maps into the frequency domain to obtain invariant features [15, 19, 20]. In some cases, Fourier transforms of the objects are used directly as the shape features [15, 19, 20, 21, 22]. Our method also uses 2D spatial maps to capture shape features. However, the shape features are based on 3D morphing and they capture both geometric and topological properties.

3. Feature Extraction and 3D Object Matching

3.1. Overview of Feature Extraction

The feature extraction procedure consists of the following steps: First, the 3D object are translated so that the object's centroid coincides with the origin of the 3D coordinate system. Next, the object is scaled so that the furthest 3D point on the object is 1 unit distance away from the centroid. Then, Principal Component Analysis (PCA) is performed on the 3D points on the object to align the major and minor axes of the object to the first and second eigenvectors of PCA. After the alignment process, the Distance Map (DM) and the Surface Penetration Map (SM) are extracted from the object.

3.2. Distance Map

After the scaling process discussed in the Section 3.1, a 3D object can fit inside a bounding sphere of unit radius, possibly with one or more points falling exactly on the sphere's surface. The object's centroid is located at the center of the sphere, which is also the origin of the 3D coordinate system. Note that the magnitude $\|\mathbf{p}_i\|$ of a point $\mathbf{p}_i$ on the object's surface is equal to the distance of the point from the center of the sphere. Furthermore, the length $d_i = 1 - \|\mathbf{p}_i\|$ is the distance that it takes to travel from the object point $\mathbf{p}_i$ to the sphere's surface along the direction of $\mathbf{p}_i$ (Fig. 1). Therefore, the sum of all d_i would correspond to the total amount of energy required to deform the object into a sphere. In other words, the values d_i measure the complexity of the shape of the object.

Instead of recording all the d_i values, an average map is computed by dividing the bounding sphere as follows: The longitudinal angle of $\theta = 0°$ to $360°$ is divided into 64 equal intervals and the latitudinal angle of $\phi = -90°$ to $+90°$ into 64 equal intervals. This process partitions the sphere into 64×64 pyramidal sections, corresponding to 64× 64 entries of the map $D(\theta, \phi)$, where θ and ϕ identify the angles at the center of the corresponding pyramidal section. Each entry (θ, ϕ) records the mean distance averaged over all d_i's of the points $\mathbf{p}_i$ contained in section (θ, ϕ). These $D(\theta, \phi)$ values form the Distance Map, which describes the geometry of the object's shape.

Note that the Distance Map can also capture information about the curvature of the object's surface. Each value $D(\theta, \phi)$ corresponds to the spherical coordinate of an average point (ρ, θ, ϕ) with $\rho = 1 - D$ in the section (θ, ϕ). The average point (ρ, θ, ϕ) is equivalent to the average of all points on all the object's surfaces within the section (θ, ϕ). The coordinates of the average points in different sections can be used to estimate the normals of the average surfaces at the average points, which can, in turn, be used to estimate the average curvature of the surfaces. Therefore, the Distance Map can be regarded as capturing all the information about a low-pass filtered version of the original object.

3.3. Surface Penetration Map

Consider a point $\mathbf{q}$ on the surface of the object. If a ray is shot from the center of the bounding sphere through $\mathbf{q}$ to the sphere's surface, the ray may penetrate one or more surfaces depending on the object's topology and concavity. Then, the mean number of surfaces $S(\theta, \phi)$ within the section (θ, ϕ) would describe the topology and concavity of the section, and the whole Surface Penetration Map would describe the topology and concavity of the entire object.

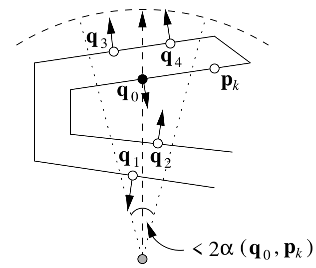

Figure 2. Counting number of object surfaces. A ray (dashed arrow) is shot from the center (gray dot) of the sphere (dashed curve) through an object point (black dot) to obtain a cone (dotted lines) with a small angle. The other object points (white dots) within the cone are considered. Checking the number of times the surface normals (arrows) at the points change direction, compared to the ray's direction, in increasing distance of the points from the sphere's center, gives the number of surfaces that the ray penetrates.

Direct computation of the intersection of a ray with surfaces that are made up of triangular meshes is quite complicated. Instead, a simpler method is developed: Consider an object point $\mathbf{q}_0$ and its corresponding ray $\mathbf{r}(\mathbf{q}_0)$ from the sphere's center to $\mathbf{q}_0$. Obtain the connected neighbors $\mathbf{p}_i$ of $\mathbf{q}_0$ from the triangular meshes that make up the object's surfaces. Among these neighbors, one of them, say $\mathbf{p}_k$, is nearest to $\mathbf{q}_0$ in terms of the angle $\alpha(\mathbf{q}_0, \mathbf{p}_k)$ between them:

$$\alpha(\mathbf{q}_0, \mathbf{p}_k) = \cos^{-1} \frac{\mathbf{q}_0 \cdot \mathbf{p}_k}{\|\mathbf{q}_0\| \, \|\mathbf{p}_k\|} \, . \tag{1}$$

Now, form a cone with its vertex located at the sphere's center, extending along the direction of $\mathbf{q}_0$, and with an angle smaller than $2\alpha(\mathbf{q}_0, \mathbf{p}_k)$ (Fig. 2). Then, all other object points $\mathbf{q}_i, i = 1, \ldots, n$, that are contained in the cone must lie on different surfaces than $\mathbf{q}_0$. Next, sort the points $\mathbf{q}_i, i = 0, \ldots, n$, in increasing order of their distances $\|\mathbf{q}_i\|$ from the sphere's center, and compute the inner product s_i between $\mathbf{q}_0$ and the surface normals $\mathbf{n}(\mathbf{q}_i)$ at the points $\mathbf{q}_i$:

$$s_i = \mathbf{q}_0 \cdot \mathbf{n}(\mathbf{q}_i) \, . \tag{2}$$

The surface normal $\mathbf{n}(\mathbf{q}_i)$ at $\mathbf{q}_i$ can be easily computed from the cross product of the vectors that connect $\mathbf{q}_i$ to its neighbors in the triangular mesh that contains $\mathbf{q}_i$. Finally, check the signs of s_i in increasing order of distance $\|\mathbf{q}_i\|$. A sign change indicates that the ray $\mathbf{r}(\mathbf{q}_0)$ has penerated another surface. So, the number of surfaces that the ray penetrates is equal to the number of sign changes plus one.

Same as the Distance Map, the bins of the Surface Penetration Map is obtained by partitioning the bounding sphere into 64×64 sections. The mean number of surfaces penetrated by the rays within a section (θ, ϕ) is computed and recorded as the bin value $S(\theta, \phi)$.

3.4. 3D Object Matching with Fourier Transform

It is mentioned in Section 3.1 that the objects are aligned using PCA before feature maps are extracted from them. However, similar objects aligned by PCA may still be mirror reflections or rotated versions of each other because the eigenvectors returned by PCA may point at opposite directions. To handle this problem, as well as possible misalignment due to the variability of the objects in the same class, Fast Fourier Transform (FFT) is performed on the feature maps. The amplitudes $\mathcal{D}(u, v)$ and $\mathcal{S}(u, v)$ of the FFTs of the feature maps $D(\theta, \phi)$ and $S(\theta, \phi)$ are invariant to rotation and reflection.

In theory, the use of Fourier transform alone can already capture features that are invariant to rotation and reflection. Nevertheless, in our implementation, PCA is still applied to align the objects. This alignment process will reduce the variation of the map entries between similar but unaligned objects due to discretization of the spatial feature maps into 64×64 entries. Increasing the spatial resolution of the feature maps may reduce the need to perform PCA but will drastically increase the storage and computational requirements of subsequent processes.

The dissimilarity between two objects are computed using weighted normalized Euclidean distance of the Fourier transforms of the feature maps:

$$d(\mathcal{D}, \mathcal{D}') = \sum_{u,v} \max\{\mathcal{D}(u, v), \mathcal{D}'(u, v)\} [\overline{\mathcal{D}}(u, v) - \overline{\mathcal{D}'}(u, v)]^2 \tag{3}$$

where $\overline{\mathcal{D}}$ and $\overline{\mathcal{D}'}$ are the normalized values of $\mathcal{D}$ and $\mathcal{D}'$. The same dissimilarity measure is defined for $d(\mathcal{S}, \mathcal{S}')$.

4. Experimental Tests

The test database consists of 34 categories of 52 sample objects (Fig. 3). Some categories have only one sample object each while other categories have more. To generate more sample objects and to assess the invariance property of our method, 13 new versions are created from each sample object by arbitrary rotation, reflection, and non-uniform scaling (see following sections for details). In total, 728 samples are created in the object database.

Three sets of tests were performed to assess various properties of our method: (1) invariance to rotation and reflection, (2) invariance to non-uniform scaling, and (3) overall retrieval performance.

4.1. Invariance to Rotation and Reflection

One sample object from each of the 34 categories was chosen. Five new versions of each sample were generated with the following transformations: 2 with aribitrary rotations, 1 with reflection about the Y-Z plane, 2 with arbitrary rotations followed by reflection about the Y-Z plane. Altogether, there were $6 \times 34 = 204$ sample objects.

Each sample object was chosen in turn as the query object, and the other objects were compared with the query object in terms of the distances $d(\mathcal{D}, \mathcal{D}')$ and $d(\mathcal{S}, \mathcal{S}')$. Average precision-recall curves were plotted for retrievals using Fourier transforms of Distance Map (DM) and Surface Penetration Map (SM). In addition, retrieval performance was also assessed for retrievals using raw feature maps DM and SM, which served as the baseline results. Weighted normalized Euclidean distance was used in all test.

Figure 4(a) shows that our method is very accurate and is invariant to arbitrary rotation and reflection when the Fourier transforms of the feature maps are used. On the other hand, retrieval using the raw feature maps are not invariant, even though the objects have been aligned using PCA, as already explained in Section 3.4.

4.2. Invariance to Non-uniform Scaling

One sample object from each of the 34 categories was chosen. Eight new versions of each sample were generated as follows: 2 with non-uniform scaling along the X-, Y-, and Z-axes, 2 with arbitrary rotations followed by non-uniform scaling, 2 with non-uniform scaling followed by reflection, and 2 with rotations followed by non-uniform scaling and reflection. For all cases, the scaling factors ranged from 0.8 to 1.2. In total, $9 \times 34 = 306$ sample objects were used in this test. Retrieval tests were performed with each sample object serving as the query object.

Figure 4(b) illustrates that our method is very accurate and is quite invariant to non-uniform scaling, coupled with rotation and/or reflection, when the Fourier transforms of the feature maps are used. Again, retrieval with the raw feature maps are not invariant.

4.3. Overall Retrieval Performance

For this test, all 52 sample objects of the 34 categories were used. Moreover, all the 14 versions of the 52 samples were generated with various combinations of arbitrary rotation, reflection, and non-uniform scaling, making up a total of 728 sample objects. Each sample object served, in turn, as a query object, and all other objects were compared with the query object and ranked according to increasing distance. Sample objects that belong to the same category as the query object were considered as relevant objects.

For comparison, histograms of morphing distance and surface penetration were generated as the features, and weighted normalized Euclidean distance was used to compare the histograms in the retrieval tests. As discussed in Section 2, many existing methods use histogram features that do not capture spatial information of the object's features. So, results of retrieval using Distance Histogram and

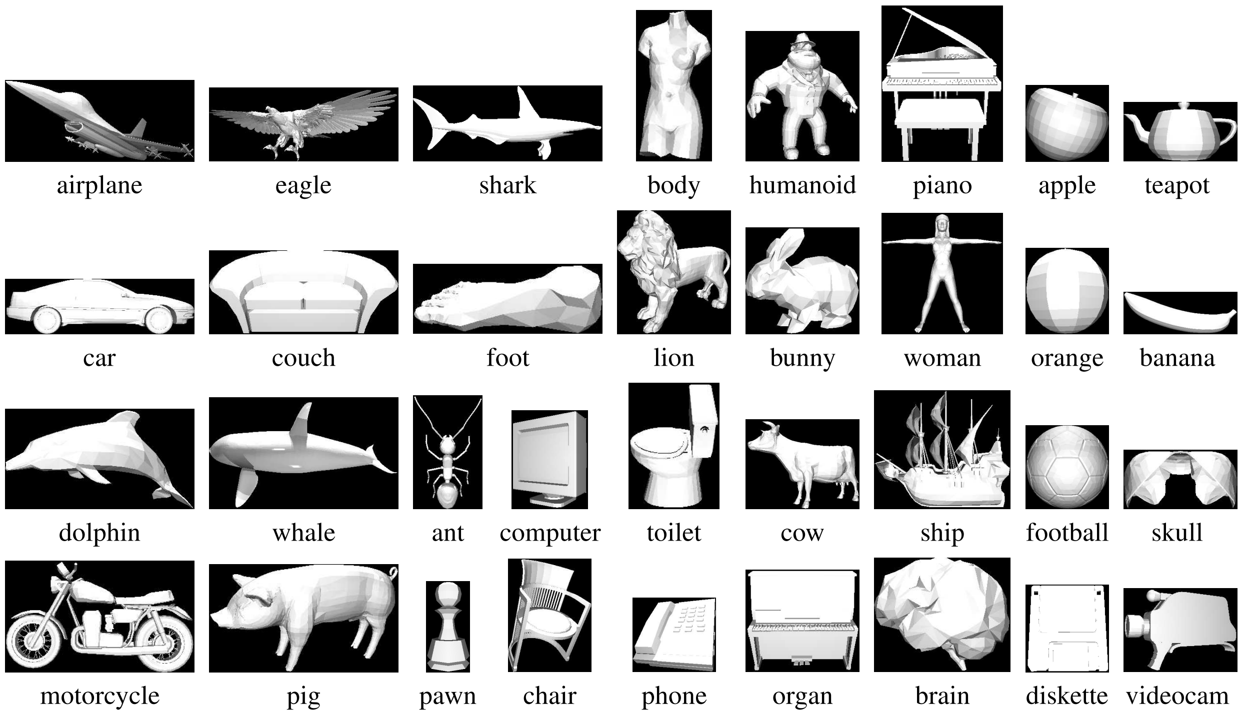

Figure 3. Samples of the 34 categories of objects.

Surface Penetration Histogram can serve as the baseline results, which represent the performance of a class of existing methods that use histogram features. In addition, to assess the accuracy of weighted normalized Euclidean distance, retrievals were also performed using Fourier transforms of feature maps coupled with normalized Euclidean distance.

Figure 5 shows that our method of 3D object retrieval using Fourier transforms of feature maps is not only invariant to linear transformation, it can also retrieve other instances of the same category very accurately. Retrieval precision remains above 0.86 even at a recall rate of 1.0. These results show that our method can perform fine discriminations of 3D shapes, such as among the different four-legged animals like lions, rabbits, cows, and pigs (Fig. 3). Retrieval using weighted normalized Euclidean distance is slightly more accurate than that using the unweighted version.

In comparison, retrieval using histograms is less accurate than that using Fourier transform of feature maps because histograms do not capture the spatial information about the object's features. Consequently, retrieval precision decreases rapidly with increasing recall rate. The precision-recall curves reported in existing work [15, 17, 18, 19, 20] show the same trend of rapid decrease in precision with increasing recall. Nevertheless, retrieval using Surface Penetration Histogram is still more accurate than retrieval using raw feature maps because histogram is invariant to linear transformation whereas raw feature maps are not invariant.

5. Conclusion

This paper presented a method of retrieving 3D models using geometric and topological spatial feature maps called Distance Map and Surface Penetration Map. The feature maps are based on the amount of effort required to morph a 3D object into a canonical sphere without performing explicit 3D morphing. Fourier transforms of the feature maps are used for comparison and retrieval. Because spatial information about the objects' features is preserved in the feature maps, retrieval performance based on the Fourier transforms of these maps is very accurate and is invariant to rotation, reflection, and scaling. Test results show that retrieval precision remains above 0.86 even for recall rate of 1.0.

Since the retrieval performance of the individual feature map is already very accurate, there is no need to combine the two feature maps in a single retrieval for the current set of testing objects. We note that with more object categories and more instances in each category, and perhaps with different similarity criteria, there may be a need to combine the geometric and topological feature maps to complement each others' strength. In our continuing research, we plan to determine the situations under which the combination of the two maps significantly outperforms the individual maps.

Acknowledgment

This research is supported by ARF R-252-000-137-112.

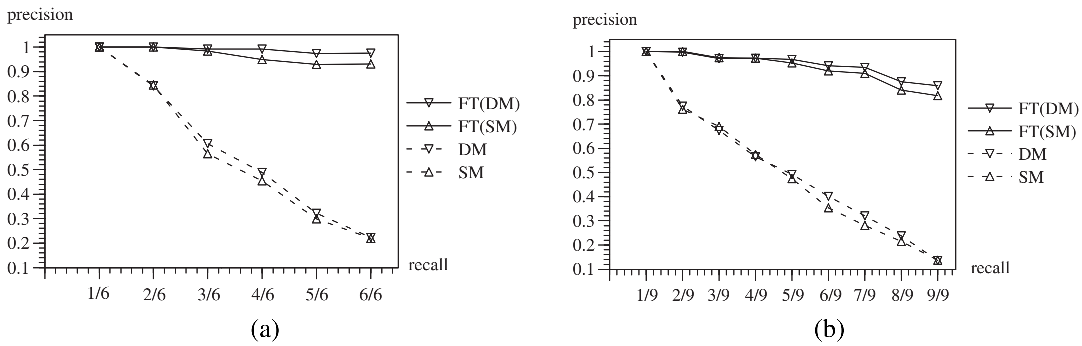

Figure 4. The retrieval results show that Fourier transforms of feature maps (FT(DM), FT(SM)) are very accurate and are invariant to (a) rotation and reflection, and (b) non-uniform scaling. In contrast, the raw feature maps (DM, SM) are not invariant.

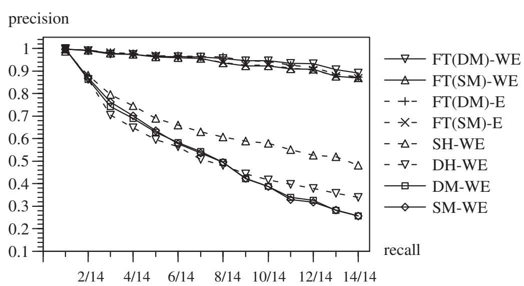

Figure 5. Comparison of retrieval performance. DM, SM: Distance and Surface Penetration Maps; FT: Fourier transform; DH, SH: Distance and Surface Penetration Histograms; WE, E: weighted and unweighted normalized Euclidean distances.

References

[1] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In *Proc. ICCV*, volume 1, pages 454–461, 2001.

[2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI*, 24(4):509–522, 2002.

[3] D. Cohen-Or, D. Levin, and A. Solomovoci. Three-dimensional distance field metamorphosis. *ACM Trans. Graphics*, 17(1):116–141, 1998.

[4] M. Elad, A. Tal, and S. Ar. Content based retrieval of VRML objects - an iterative and interactive approach. In *Proc. 6th Eurographics Workshop on Multimedia*, 2001.

[5] M. Hebert, K. Ikeuchi, and H. Delingette. A spherical representation for recognition of free-form surfaces. *IEEE Trans. PAMI*, 17(7):681–690, 1995.

[6] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proc. SIGGRAPH*, 2001.

[7] H.-P. Kriegel, T. Schmidt, and T. Seidl. 3D similarity search by shape approximation. In *Proc. 5th Int. Symposium on Large Spatial Databases*, volume 1262, pages 11–28, 1997.

[8] H.-P. Kriegel and T. Seidl. Approximation-based similarity search for 3-D surface segments. *GeoInformatica Journal*, 2:113–147, 1998.

[9] G. Mori, S. Belongie, and J. Malik. Shape contexts enable efficient retrieval of similar shapes. In *Proc. CVPR*, 2001.

[10] M. Novotni and R. Klein. A geometric approach to 3D object comparison. In *Proc. Int. Conf. on Shape Modeling and Applications*, pages 167–175, 2001.

[11] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Matching 3D models with shape distribution. In *Proc. Shape Modeling International*, 2001.

[12] E. Paquet and M. Rioux. Nefertiti: A query by content system for three-dimensional model and image databases management. *Image and Vision Computing*, 17:157–166, 1999.

[13] E. Paquet, M. Rioux, A. Murching, T. Naveen, and A. Tabatabai. Description of shape information for 2-D and 3-D objects. *Signal Processing: Image Communication*, 16:103–122, 2000.

[14] S. Ruiz-Correa, L. G. Shapiro, and M. Melia. A new signature-based method for efficient 3D object recognition. In *Proc. CVPR*, 2001.

[15] D. Saupe and D. V. Vranic. 3D model retrieval with spherical harmonics and moments. In *Proc. 23rd DAGM Symposium (LNCS 2191)*, pages 392–397, 2001.

[16] H.-Y. Shum, M. Herbert, and K. Ikeuchi. On 3D shape similarity. In *Proc. CVPR*, pages 526 – 531, 1996.

[17] M. T. Suzuki, T. Kato, and N. Otsu. A similarity retrieval of 3D polygonal models using rotation invariant shape descriptors. In *Proc. IEEE Conf. SMC*, pages 2946–2952, 2000.

[18] M. T. Suzuki, T. Kato, and H. Tsukune. 3D object retrieval based on subjective measures. In *9th Int. Conf. on Database and Expert Systems Applications*, 1998.

[19] D. V. Vranic and D. Saupe. 3D shape descriptor based on 3D fourier transform. In *Proc. of the EURASIP Conf. on Digital Signal Processing for Multimedia Communications and Services*, pages 271–274, 2001.

[20] D. V. Vranic, D. Saupe, and J. Richter. Tools for 3D object retrieval: Karhunen-Loeve transform and spherical harmonics. In *Proc. of IEEE Workshop on Multimedia Signal Processings*, pages 293–298, 2001.

[21] C. Zhang and T. Chen. Efficient feature extraction for 2D/3D objects in mesh representation. In *Proc. ICIP*, 2001.

[22] C. Zhang and T. Chen. Indexing and retrieval of 3D models aided by active learning. In *Proc. ACM Multimedia*, 2001.