3D model retrieval based on Grid Sphere and Dodecahedral Silhouette Descriptors

Jau-Ling Shih* and Hong-Yu Chen

Department of Computer Science and Information Engineering,
Chung Hua University, Hsinchu, Taiwan, R.O.C
E-mail address: sjl@chu.edu.tw

Abstract

With the development of computer graphics and virtual realities, the demand for a content-based 3D model retrieval system becomes urgent. In this study, two features, grid sphere and dodecahedral silhouettes, are proposed and combined for 3D model retrieval. The experiments are conducted on a 3D model database. Experiment results show that the proposed methods are superior to others.

Keywords: 3D model retrieval, grid sphere descriptor, dodecahedral silhouette descriptor

1. Introduction

In recent year, with the development of computer graphics and virtual realities, the management of 3D models becomes an important issue. Thus, developing an automatic and user-friendly content-based 3D model retrieval system become urgent.

The major challenge for such a system is how to extract proper features to represent 3D models and fast search similar models using these features.

An approach is to extract 3D features directly from the 3D shape. Saupe et al. [1] use spherical harmonics to obtain a multi-resolution representation of 3D models. However, this method requires a priori registration based on the principal component analysis (PCA). However, in our experiences, some similar 3D models exist different principal axes. Funkhouser et al. [2] develop a 3D model search engine. The feature of adaptive spherical harmonics is used to efficiently measure similarity without aligning model's orientation by PCA.

Some features to represent the 3D models are based on the histograms of geometric statistics. Shape histograms [3] characterize the area of intersection of a 3D model with a collection of concentric shells and sectors. Quadratic form distance measure is employed to compute the distance between the histogram bins. The MPEG-7 shape spectrum descriptor (SSD) [4] calculates the histogram of the curvatures of all points on the 3D surface. SSD represents the distribution of geometric characteristcs and is robust to tessellation of 3D polygonal models. Osada et al. [5] propose five features, A3, D1, D2, D3, and D4, to represent 3D models with probability distributions of geometric properties computed from randomly selected points on the model's surface. For instance, D2, the best feature among these five features, is the distribution of distances between two random points.

The simplest way is to describe a 3D object by its 2D silhouettes from different views [6]. Users can find similar 3D models by 2D shape features. Several methods [2, 7] provide a 2D query interface to facilitate 3D model retrieval based on 2D silhouettes. Super and Lu [8] exploit 2D silhouette contours for 3D object recognition. Curvature and contour scale space are extracted to represent each silhouette. Chen et al. [6] introduced the LightField descriptor to represent 3D models. The LightField descriptor is computed from 10 silhouettes rendered from vertices of dodecahedron over a hemisphere. Each silhouette is represented by a 2D binary image. The Zernike moments and Fourier descriptors are employed to describe each binary image. Since a 3D model may be rotated or deformed, the number of 2D silhouettes must be large enough to represent a 3D model. Thus, a set of 10 LightField descriptors is applied to enhance the robustness. That is, 100 (10×10) silhouettes are used to represent a 3D model. However, if the number of silhouettes increases, the retrieval speed will be greatly deceased. Thus, choosing the appropriate number of silhouettes to represent 3D models is an important issue. On the other hand, the silhouettes can only represent the exterior shape information of a 3D model. The description for 3D model interior is short. Therefore, in this study, two features, grid sphere descriptor and dodecahedral silhouettes, are proposed and combined to retrieve 3D models. The grid sphere descriptor can extract the 3D model interior features and the dodecahedral silhouettes can represent the exterior features.

The rest of the paper is organized as follows. In Section 2, the proposed feature is introduced. Section 3 gives the experimental results to show the

effectiveness of the proposed features. Finally, conclusions is given in Section 4.

2. Feature extraction

Two features, grid sphere and dodecahedral silhouettes are proposed in this study for 3D model retrieval.

2.1. Grid Sphere Descriptor(GSD)

The grid sphere descriptor(GSD) is introduced in this subsection to describe the shape information within a 3D model. The main steps for computing the (GSD) and measuring the similarity are described as following:

(1) First, the 3D model is decomposed into a $64 \times 64 \times 64$ voxel grid. A voxel is regarded as valid if there is a polygonal surface located within this voxel and invalid otherwise. To be invariant to translation and scaling, the object's mass center is moved to the location (32, 32, 32) and the average distance from valid voxels to the mass center is scaled to be 16, which provides adequate resolution for discriminating objects while filtering out those high-frequency polygonal surfaces in the complex components of a 3D model. [9]

(2) After the mass center being found, the normalized 3D model is segmented by a $32 \times 64 \times 64$ sphere grid again (see Fig. 1(a)). There are 32 shells and each shell is segmented by 64×64 grid. Fig. 1(b) is a 2D profile on sphere grid. Each grid in a 3D model is shown as Fig. 1(c). A sphere grid is regarded as valid if there is a polygonal surface located within it and invalid otherwise.
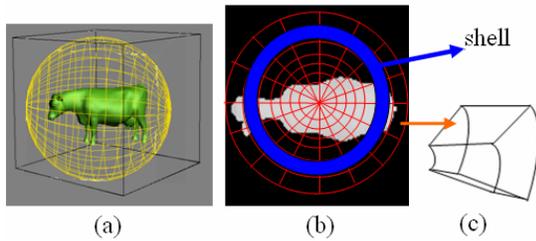

Fig. 1: the sphere decomposed 3D cow model. (a) The 3D model is segmented by sphere grid. (b) 2D profile on sphere grids. (c) Each sphere grid.

(3) For each shell, the number of valid grids is calculated to get the GSD. The GSD $\mathbf{x}^G$ is defined as: $\mathbf{x}^G = [x(1), x(2), \ldots, x(32)]^T$, where $x(i)$ is the total number of valid grids in the i-th shell.

Let $\mathbf{x}^G$ and $\mathbf{y}^G$ denote the grid sphere descriptors of query model q and matching model s, respectively. The distance between $\mathbf{x}^G$ and $\mathbf{y}^G$ is defined as:

$$Dis_{q,s}^G = \sum_{i=1}^{32} \left| x^G(i) - y^G(i) \right|.$$

2.2. Dodecahedral Silhouette Descriptor (DSD)

In this subsection, the dodecahedral silhouette descriptor is described to represent the outside information of a 3D model. Since a 3D model may be rotated or deformed, the number of 2D silhouettes must be large enough to represent a 3D model. However, if the number of silhouettes increases, the retrieval speed will be greatly deceased. Thus, we want to choose the appropriate number of silhouettes to represent 3D models is an important issue. Fig. 2(a) is a diagram dodecahedron, and Fig. 2(b) show 2D silhouettes from the face of dodecahedron over a hemisphere.
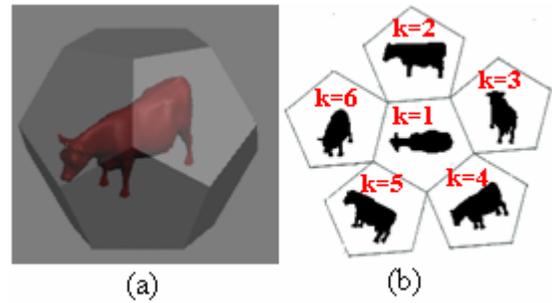

Fig. 2: (a) the 3D cow model within a dodecahedron. (b) 6 silhouettes from the face of dodecahedron over a hemisphere.

The main steps for computing dodecahedral silhouette descriptor (DSD) is described as by the following steps:

(1) First, 6 silhouettes are rendered from faces of dodecahedron over a hemisphere to represent a 3D model. Each silhouette is represented by a 2D binary image.

(2) Next, The angular radial transformation (ART) is extracted for each silhouette. The ART is the MPEG-7 region-based shape descriptor. It is defined on a unit disk in polar coordinates. From each silhouette, a set of ART coefficients F_{nm} is defined as:

$$F_{nm} = \langle V_{nm}(\rho, \theta), f(\rho, \theta) \rangle$$
$$= \int_0^{2\pi} \int_0^1 V_{nm}^{\bullet}(\rho, \theta) f(\rho, \theta) \rho \, d\rho \, d\theta,$$

where $f(\rho, \theta)$ is an image intensity function in polar coordinates and $V_{nm}(\rho, \theta)$ is the ART basis function of order n and m. The basis functions that are separable along the angular and radial directions are defined as:

$$V_{nm}(\rho, \theta) = A_m(\theta)R_n(\rho).$$

The angular and radial basis functions are defined as:
$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta) \text{ and}$$

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2\cos(\pi n\rho) & n \neq 0 \end{cases}$$

In this study, the ART has 36 coefficients ($n = 10$ and $m = 10$).

The DSD $\mathbf{x}^D$ is defined as:
$$\mathbf{x}^D = \left[(\mathbf{x}_1^D)^T, (\mathbf{x}_2^D)^T, \ldots, (\mathbf{x}_6^D)^T \right]^T,$$

where
$$\mathbf{x}_k^D = \left[x_k^D(1), x_k^D(2), \ldots, x_k^D(36) \right]^T,$$

and k is the index of silhouettes (see Fig. 2(b)), and $(\mathbf{x}_1^D)^T, (\mathbf{x}_2^D)^T, \ldots, (\mathbf{x}_6^D)^T$ are the 36 coefficient from ART.

(3) Let $\mathbf{x}^D = \left[(\mathbf{x}_1^D)^T, (\mathbf{x}_2^D)^T, \ldots, (\mathbf{x}_6^D)^T \right]^T$ and
$$\mathbf{y}^D = \left[(\mathbf{y}_1^D)^T, (\mathbf{y}_2^D)^T, \ldots, (\mathbf{y}_6^D)^T \right]^T$$

denote the DSD of q and s, respectively. For the matching operation corresponding to the i-th permutation p_i, $1 \leq i \leq 30$, the distance between $\mathbf{x}^D$ and $\mathbf{y}^D$ is defined as (see Table 1):

Table 1 30 permutations for DSD matching between a query model and a matching model.

k	1	2	3	4	5	6	k	1	2	3	4	5	6
$p_1(k)$	1	2	3	4	5	6	$p_{16}(k)$	4	1	3	6	2	5
$p_2(k)$	1	3	4	5	6	2	$p_{17}(k)$	4	3	6	2	5	1
$p_3(k)$	1	4	5	6	2	3	$p_{18}(k)$	4	6	2	5	1	3
$p_4(k)$	1	5	6	2	3	4	$p_{19}(k)$	4	2	5	1	3	6
$p_5(k)$	1	6	2	3	4	5	$p_{20}(k)$	4	5	1	3	6	2
$p_6(k)$	2	1	6	4	5	3	$p_{21}(k)$	5	1	4	2	3	6
$p_7(k)$	2	6	4	5	3	1	$p_{22}(k)$	5	4	2	3	6	1
$p_8(k)$	2	4	5	3	1	6	$p_{23}(k)$	5	2	3	6	1	4
$p_9(k)$	2	5	3	1	6	4	$p_{24}(k)$	5	3	6	1	4	2
$p_{10}(k)$	2	3	1	6	4	5	$p_{25}(k)$	5	6	1	4	2	3
$p_{11}(k)$	3	1	2	5	6	4	$p_{26}(k)$	6	1	5	3	4	2
$p_{12}(k)$	3	2	5	6	4	1	$p_{27}(k)$	6	5	3	4	2	1
$p_{13}(k)$	3	5	6	4	1	2	$p_{28}(k)$	6	3	4	2	1	5
$p_{14}(k)$	3	6	4	1	2	5	$p_{29}(k)$	6	4	2	1	5	3
$p_{15}(k)$	3	4	1	2	5	6	$p_{30}(k)$	6	5	1	2	3	4

$$Dis_{q,s}^{D,i} = \sum_{k=1}^{6} \left\| \mathbf{x}_k^D - \mathbf{y}_{P_i(k)}^D \right\|_1 = \sum_{k=1}^{6}\sum_{r=1}^{36} \left| x_k^D(r) - y_{p_i(k)}^D(r) \right|,$$

where $p_i(k)$ denotes the k-th value for the i-th permutation, $1 \leq k \leq 6$. The distance between the query model q and the matching model s is defined as:

$$Dis_{q,s}^D = \min_{1 \leq i \leq 30} Dis_{q,s}^{D,i}.$$

2.3. The combination of GSD and DSD

In order to exactly describe a the 3D model, we combine GSD and DSD, to represent the external and internal information of a 3D model. The distance on GSD as the weight of the distance on the DSD. The total distance is defined as:

$$Dis_{q,s}^T = Dis_{q,s}^G \times Dis_{q,s}^D.$$

Based on $Dis_{q,s}^T$, we can find 3D models similar to the query one by taking those with low values.

3. Experimental results

To demonstrate the effectiveness of the proposed methods for different kinds of 3D models, experiments have been conducted on a test database, and are implemented to compare the retrieval results. The performance is measured by the *recall*. The *recall* is defined :

$$recall = N / T,$$

where N is the number of relevant models retrieved and T is the total number of relevant models in the database.

Database is established to test the performance of invariance to deformations. To derive Database, 20 models are first selected as the seed models (see Fig. 3). Then, each seed model is deformed by 24 kinds of transformations, including 4 geometric deformations, 2 scalings, 13 rotations, and 5 various resolutions (see Fig. 4). Thus, there are totally 500 models in Database.

The comparison of *recall* and time of DSD (using 6 silhouettes) and LightField (using 10 silhouettes) is shown in Table 2. We can see that, the recall values of two methods are almost equivalence, but DSD is faster than LightField on the retrieval time (improvement 0.32%). As shown in Table 3, the *recall* of GSD is 0.8270 and only 0.149 sec is needed to retrieve the similar 3D models. Moreover, the combination of DSD and GSD (DSD+GSD) has the similar high recall value as the combination of LightField and GSD (LightField+GSD). Note that, DSD+GSD is still more faster than LightField+GSD (improvement 0.30%). Finally, our method, DSD+GSD, is compared with

other methods, D2[5], spherical harmonics (SH)[1], and shape spectrum descriptor (SSD)[4] (see Table 4). Although other methods can fast retrieve 3D models, our method has the greatly higher *recall* (over 90%) than others.
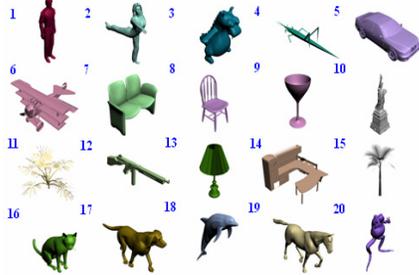
Fig. 3 : 20 seed models in the test database.

Fig. 4 : 25 deformed models for the dragon class.

Table 2 Recall and time of DSD and LightField [6].

	Recall	Time (sec)
DSD	**0.8290**	**0.5600**
LightField [6]	0.8272	0.8342

Table 3 Recall and time of GSD and the combined descriptors.

	Recall	Time (sec)
GSD	0.8270	0.1489
DSD +GSD	**0.9197**	**0.6512**
LightField [6]+GSD	0.9263	0.9235

Table 4 The comparison of recall and time of the proposed method and others.

	Recall	Time (sec)
DSD +GSD	**0.9197**	**0.6512**
D2[5]	0.5623	0.2233
SH[1]	0.4875	0.5658
SSD[4]	0.7180	0.3846

4. Conclusion

With the development of computer graphics and virtual realities, the demand for a content-based 3D model retrieval system becomes an important issue. In this study, two features, grid sphere descriptor (GSD) and dodecahedral silhouettes descriptor (DSD), are proposed and combined for 3D model retrieval. The experiments have been conducted on a 3D model database. Experiment results show that the proposed methods are superior to others.

5. Acknowledgments

6. Reference

[1] D.V. Vranic, D. Saupe, and J. RICHTER, "Tools for 3D-object retrieval: Karhunen-Loevetransform and spherical harmonics", Proc. IEEE Workshop Multimedia Signal Processing, pp. 293–298, 2001.

[2] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, "A Search Engine for 3D Models", ACM Trans. on Graphics, Vol. 22 No. 1, pp. 83-105, 2003.

[3] M. Ankerst, G. Kastenmuller, H.P. Kriegel, and T. Seidl, "3D shape histograms for similarity search and classification in spatial databases", Proc. Int. Symposium on Spatial Databases (SSD), Hong Kong, China, pp. 207-226, 1999.

[4] S. Manjunath, P. Salembier, and T. Sikora, "Introduction to MPEG-7 Multimedia Content Description Interface", John Wiley & Sons Ltd., pp. 247-260, 2002.

[5] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape Distributions", ACM Trans. on Graphics, Vol. 21 No. 4, pp. 807-832, 2002.

[6] D.Y. Chen, X.P. Tian, Y.T. Shen, M. Ouhyoung, "On visual similarity based 3D model retrieval", Computer Graphics Forum, Vol. 22 No. 3, pp. 223-232, 2003.

[7] J. Loffler, "Content-based retrieval of 3D models in distributed web databases by visual shape information", Proceedings of Information Visualization, pp. 82-87, 2000.

[8] B.J. Super, H. Lu, "Evaluation of a hypothesizer for silhouette-based 3-D object recognition", Pattern Recognition 36, pp. 69-78, 2003.

[9] J. L. Shih, C. H. Lee, and J. T. Wang, "3D Object Retrieval System Based on Grid D2", Electronics Letters Vol. 41 No. 4, pp. 23-24, 2005.