3D Object Retrieval System Based on Grid D2

Jau-Ling Shih, Chang-Hsing Lee, and Jian Tang Wang

With the increase of the number of 3D objects available on the digital libraries, the demand for a content-based 3D object retrieval system becomes urgent. In this paper, we propose a novel feature, grid D2, for 3D object retrieval. Experiment results show that the proposed method is superior to others.

Introduction: With the development of computer graphics and virtual realities, 3D objects will be as prevalent as other multimedia data in the future. Thus, developing an automatic content-based 3D object retrieval system is necessary. The major problem for such a system is how to extract proper features to represent variable shapes in a 3D model and effectively search similar 3D objects using these features. The simplest way is to represent a 3D object by its 2D silhouettes from different views [1], users can find similar 3D objects by a set of 2D shape features extracted from these 2D silhouettes. Since a 3D object could be rotated or deformed, the number of 2D silhouettes must be large enough to represent a 3D model. On the other hand, if the number of silhouettes increases, the retrieval speed will be deceased. In addition, even many silhouettes are exploited, if a complex 3D model is rotated slightly, the extracted features do not guarantee to be invariant.

A better approach is to extract 3D features directly from a 3D object. Before extracting the 3D feature vectors, some methods [2-3] need to register the 3D models based on the principal
component analysis (PCA). However, we find that some similar 3D objects have different principal axes. As shown in Fig. 1, small differences exist in the handles of similar mugs could affect the principal axes significantly. Saupe et al. [4] have used spherical harmonics to obtain a multi-resolution representation of 3D models. However, this method also requires a priori registration with principal axes. Recently, Funkhouser et al. [5] developes a 3D model search engine, which used adaptive spherical harmonics to compute discriminating similarity measures without repairing degenerate objects or aligning object’s orientation. Some other features to represent the 3D objects are based on the histograms of geometric statistics. Ankerst et al. [6] proposes shape histograms to characterize the area of intersection with a collection of concentric spheres. The shape spectrum descriptor (SSD) [7] used in MPEG-7 calculates the histogram of the curvatures of all points on the 3D surface. The advantages of SSD are that it can represent the distribution of geometric characteristics without aligning the objects and is robust to tessellation of 3D polygonal models. Osada et al. [8] propose a set of features to represent 3D objects with probability distributions of geometric properties computed from randomly selected points on the object’s surface. Five features, including A3, D1, D2, D3, and D4, are proposed for measuring shape distributions. For instance, D2, the best feature among these five features, is the distribution of distances between two random points. These features are invariant to tessellation of 3D polygonal models, since points are randomly selected from the object’s surface. However, they are sensitive to small deformation due to noise, cracks, or insertion/removal of polygons, since sampling is area weighted. As shown in Fig. 2, to finely represent the complex components of a 3D object (for example, the propeller and wheels of the aircraft), a 3D model often requires many polygons. The random sampling of a 3D model would be dominated by those complex
components. Thus, a novel feature, called grid D2, will be proposed to improve the performance of the traditional D2.

**Feature extraction:** First, the 3D model is decomposed by a voxel grid. A voxel is regarded as valid if there is a polygonal surface located within this voxel and invalid otherwise. Then the distribution of distances between two valid voxels instead of two points on the surface are calculated. Therefore, the area weighted defect in the sampling process will be greatly reduced since each valid voxel is weighted equally irrespective of how many points located within this voxel. The main steps for computing the grid D2 are described as follows:

1. First, a 3D model is segmented into a $2R \times 2R \times 2R$ voxel grid (see Fig. 3). To be invariant to translation and scaling, the object’s mass center is moved to the location $(R, R, R)$ and the average distance from valid voxels to the mass center is scaled to be $R/2$. $R$ is set as 32, which provides adequate resolution for discriminating objects while filtering out those high-frequency polygonal surfaces in the complex components of a 3D object.

2. Two valid voxels are randomly selected and their distance is measured. A total of $M$ distances are evaluated from the set of valid voxels. A histogram containing 256 bins is constructed: $H = \{V_1, V_2, \ldots, V_{256}\}$, where $V_i$ denotes the number of distances within the range of the $i$-th bin. In order to normalize the distribution, the grid D2 (GD2) is defined as:

$$GD2 = \left[ \frac{V_1}{M}, \frac{V_2}{M}, \frac{V_3}{M}, \ldots, \frac{V_{256}}{M} \right]$$

where $M$ is set as $64^3$. As shown in Fig. 4, we can see that the D2 distributions are clearly different while GD2 distributions are similar for these two similar airplanes.
**3D Object Retrieval:** Given a query 3D object $q$ and any matching 3D object $s$, the GD2 difference between $q$ and $s$ is defined as:

$$\text{Dis}_{GD2_{q,s}} = \sum_{i=1}^{256} |GD2_i^q - GD2_i^s|.$$  

Then, the similarity measure between $q$ and $s$ is defined as:

$$\text{Sim}_{GD2_{q,s}} = \frac{1}{\text{Dis}_{GD2_{q,s}}}.$$  

Note that the larger $\text{Sim}_{GD2_{q,s}}$ is, the more similar a matching object is to the query one. Based on $\text{Sim}_{GD2_{q,s}}$, we can find objects similar to the query one by taking those with high values.

**Experimental Results:** To demonstrate the effectiveness of the proposed method on different kinds of 3D objects, experiments have been conducted based on the Princeton Shape Benchmark test database [9]. The Princeton Shape Benchmark provides a database containing 1814 3D objects (161 classes) for evaluating shape-based retrieval and analysis algorithms. The performance is measured by the recall which is defined as:

$$\text{recall} = \frac{N}{T},$$

where $N$ is the number of relevant objects retrieved and $T$ is the total number of relevant objects. To show the effectiveness of GD2, the performances of GD2, D2, spherical harmonics (SH), and shape spectrum descriptor (SSD) are compared. The experimental result is shown in Table 1. We can see that the proposed GD2 outperforms other methods.
Conclusions: In this paper, a new feature, grid D2 (GD2), is proposed for automatic 3D object retrieval. The main research contribution is that the new shape descriptor is both discriminating and robust. Experimental results also show that the proposed method is superior to others.

Acknowledgment: This research was supported in part by the National Science Council, Taiwan under Contract NSC 93-2213-E-216-014 and Chung Hua University, Taiwan under Contract CHU-93-TR-004.

Reference:


Authors’affiliations:
Jau-Ling Shih, Chang-Hsing Lee, and Jian Tang Wang (Department of Computer Science and Information Engineering, Chung Hua University, No. 707, Sec.2, WuFu Rd., Hsinchu, Taiwan, R.O.C). The corresponding author is Jau-Ling Shih (sjl@chu.edu.tw).

Table captions:
Table 1. The recall for the test database

Figure captions:
Fig. 1 The similar mugs with different principal axes
Fig. 2 The 3D model of aircraft
Fig. 3 Diagram of voxel grid for a 3D model
Fig. 4 D2 and GD2 distributions for two similar airplane objects
Table 1

<table>
<thead>
<tr>
<th>Feature</th>
<th>recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2</td>
<td>0.175</td>
</tr>
<tr>
<td>SSD</td>
<td>0.201</td>
</tr>
<tr>
<td>SH</td>
<td>0.245</td>
</tr>
<tr>
<td>GD2</td>
<td>0.257</td>
</tr>
</tbody>
</table>
Figure 1
Figure 3
Figure 4