2011 International Conference on Asia Agriculture and Animal
IPCBEE vol.13(2011) © (2011)IACSIT Press, Singapore

View Based Rotation Invariant 3D Shape Retrieval

Chi-Chou Kao+

Department of Computer Science and Information Engineering
National University of Tainan
Tainan, Taiwan

Abstract. It is necessary to develop efficient methods for retrieving the 3D models in a large database. In the past, 3D model uses only the relationship between multi-points but losses the geometric structure of the original model. In this paper, a system, which can preserve the local geometry information of model, is presented for robust shape retrieval from 3D models. The 3D model is first represented by 2D shapes of various different angles. Each 2D shape is represented by a discrete set of *n* points. Next, an efficient algorithm for rotation invariance is proposed. In the algorithm, the rotated shape contexts together are clustered and label each cluster so that the shape contexts in each cluster have the same label. Using the histogram of label frequencies can quickly and efficiently search for similar or rotational shapes. The experimental results have shown that our system is effective and has better retrieval results than the existing systems.

Keywords: retrieval, 3D, rotation-invariance, shape based, and label.

1. Introduction

In recent years, there have been a large number of advances in 3D model acquisition technology [1-5]. Since objects can be described by various features, we evaluate the similarity by the image properties. However, the complexity is increased by the number of features. It is important to utilize a few features that are representative. To retrieve similar images, we should compute the similarity between images. Shape is a distinguishable feature which can make coarse discriminations quickly. Therefore, adding shape matching can increase the performance of an image retrieval system [6]. Accordingly, the retrieval of 3D shapes [7] is very important for indexing image or textual data in practice and research interests and we will focus on shape matching for region-based image retrieval in the paper.

Given a database of 3D shapes and a query shape, a shape retrieval algorithm is to return shapes, which is visual similarity to the query shape, from the database. Many retrieval algorithms have been proposed [8-12]. Generally, each shape is characterized by a shape descriptor. Commonly used quality criteria for shape descriptors include invariance to rigid-body transformations, scaling, bending and moderate stretching, robustness against noise and data degeneracy, and storage and computational costs. Most state-of-the-art descriptors are designed to be invariant to only rigid-body transformations and uniform scaling. Hence, it is no surprise that they do not perform well when applied to shapes having non-rigid transformations such as bending or stretching, which are obviously harder to handle due to their non-linearity and increased degrees of freedom.

The paper will present an efficient algorithm for 3D shape-based model retrieval system. The 3D model is first represented by 2D shapes of various different angles. Next, an algorithm takes into account the shape context for rotating objects when performing clustering, and assign the same shape labels to a shape and its rotation. Thus, our system can find similar shapes in the database for a query image no matter they are rotations with each other. The experimental results have shown that our system is effective and has better

+ Corresponding author. Tel.: + 886-6-2606123 # 7700; fax: + 886-6-2606125.
E-mail address: cckao@mail.nutn.edu.tw

This work was supported by National Science Council, Taiwan 99-2221-E-024-008

retrieval performance than the existing systems. The remainder of this paper is organized as follows. In section II, we outline the proposed 3D model retrieval system. We describe how to represent 3D model by 2D shapes of various different angles and an efficient rotation invariance algorithm in section III. The experimental results are summarized in section IV. Finally, we conclude in section V.

2. View Based 3D Shape Retrieval

The proposed 3D shape retrieval framework is shown in Figure 1. To retrieve efficiently 3D models, we present a view based method to transform 3D model into 2D shapes of various different angles. To quickly and efficiently search for 3D models, an efficient rotation invariance algorithm, which will be described in detail in the next section, is proposed before the feature matching step.

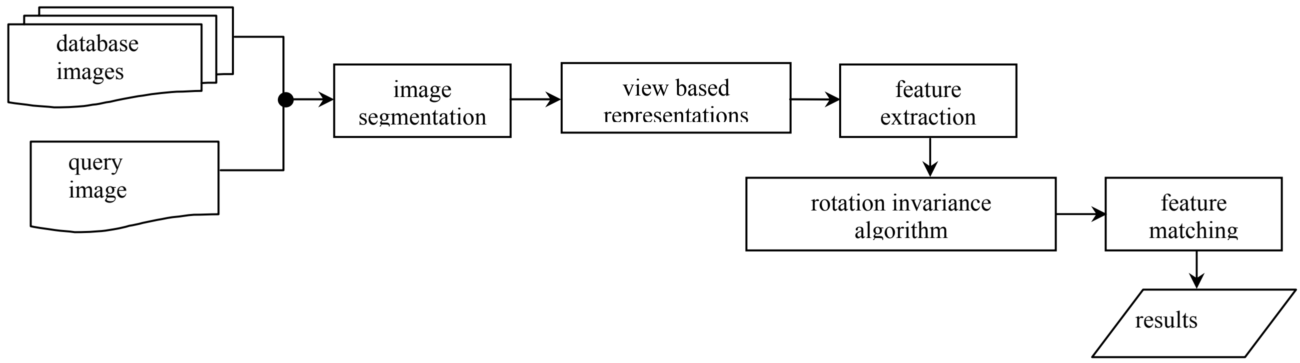

Fig. 1: 3D shape retrieval framework.

3D models look different when viewed from different viewpoints. Based on this idea, the space of views can be partitioned into view classes. Within each class, the views share a certain property. If two models are similar, they should look similar from all viewing angles. Our 3D model retrieval system is based on the view based representations that similar 3D models look similar from the same viewpoints. Therefore, a number of views (2D projections) of objects could be used to represent the models. The remainder of this section will describe the methods in which a number of views of the 3D models are generated for the purposes of similarity matching. Let *azimuth* be a polar angle in the *x*-*y* plane with positive angles indicating counterclockwise rotation of the viewpoint and *elevation* be the positive angle or negative angle of the *x*-*z* plane. According to the definitions, the azimuth ranges between 0° and 360° and the elevation ranges between -90° and 90°. Given a 3D model, we take out the 2D shapes from azimuth and elevation per 30°. Hence, we will generate 12×6=72 2D shapes from the 3D model.

To retrieve similar models, we should compute the similarity between models. The similarity is evaluated by 72 2D shapes. Let *a*, *b* be two 3D models to be compared. The similarity measure is defined as follows:

$$S(a,b) = \min \sum_{i=1}^{6} \sum_{k=1}^{12} d(I_{aik}, I_{bik})$$

where I_{aik}, I_{bik} are the 2D shapes generated from the view based representation and the distance *d* is a Euclidean distance function that represents the shortest distance along each axis between two points.

The shape rotation has significant influence on the similarity measure and retrieval performance for 2D shape retrieval from 3D models. In the next section, we will present an efficient algorithm for rotation invariance. In the algorithm, the rotated shape contexts together are first clustered and then label each cluster so that the shape contexts in each cluster have the same label. Using the histogram of label frequencies can quickly and efficiently search for similar or rotational shapes.

3. Rotation-Invariant Algorithm

In this section, we first describe retrieval by shape contexts to formulate the rotated object matching problem. Next, the rotation invariance algorithm based on the *k*-means clustering method [13] is proposed.

3.1. Problem Formulation

The shape context analysis begins by taking n samples from the edge elements on the shape. Then, a shape is represented by a discrete set of points sampled from its contours. These points can be obtained as locations of edge pixels as found by an edge detector, giving us a set $P = \{p_1, p_2, \ldots, p_n\}$, $p_i \in \boldsymbol{R}^2$, of n points. When we consider the set of vectors originating from a point to all other sample points on a shape, these n-1 vectors express the configuration of the entire shape relative to the reference point. One compact way to capture this information is the distribution of the relative positions of the remaining n-1 points in a spatial histogram. Concretely, for a point p_i on the shape, we compute a coarse histogram h_i of the relative coordinates of the remaining n-1 points,

$$h_i^k = \#\{q \neq p_i : \{q - p_i\} \in bin(m)\}.$$

This histogram is defined to be the shape context of p_i. Thus, each shape can obtain n shape contexts. To make the descriptor more sensitive to positions of nearby sample points than to those of points farther away, we use bins that are uniform in log-polar space. It should be noticed that in the absence of background clutter, the shape context of a point on a shape is made scale invariant by normalizing all radial distances by the mean distance α between the n^2 point pairs in the shape.

We observe that shape contexts will be different for different points on a single shape S; however, corresponding (homologous) points on similar shapes S and S' will tend to have similar shape contexts. By using shape contexts as shape descriptor, we can quickly determine which shapes in the database are similar to the query shape. The basic idea is using vector quantization on the shape contexts. With $|S|$ known shapes, and shape contexts computed at n sample points on these shapes, the full set of shape contexts for the known shapes consists of $|S| \times n$ d-dimensional vectors where d is the total number of bins in a shape context histogram. A compression technique for dealing with such a large amount of data is vector quantization. Vector quantization involves clustering the vectors and then representing each vector by the index of the cluster that it belongs to. These indexes are called *shape labels*.

To represent each shape with shape labels, all of the shape contexts from the known set are first transformed to d-dimensional vectors and considered as points in a d-dimensional space. These vectors are called *shape context vectors*. Then the k-means clustering can be performed to obtain k clusters and label each group by an integer in $\{1, 2, \ldots, k\}$. Each d-dimensional shape context vector is quantized to its nearest clusters and replaced by its shape label. A known shape is then simplified into a histogram of shape label frequencies. We have reduced each collection of n shape contexts (d bin histograms) to a single histogram with k bins.

Since the corresponding points on similar shapes have similar shape contexts, similar shapes will have similar shape contexts. If we use a shape label to replace a shape context, then similar shapes will have similar histogram of shape label frequencies. This property is used to match a query shape. The same vector quantization and histogram creation operation is performed on the shape contexts from the query shape. We then find nearest neighbors in the space of histogram of shape label frequencies.

However, if a shape is rotated, the distribution of shape labels will be different. The corresponding points (which are marked) have different shape contexts for one shape and its rotation by 180° and are replaced with different shape labels. Obviously, this causes dissimilar histograms of shape label frequencies between the shape and its rotation. If we perform image retrieval by using the set of shape labels directly, the rotated object will not be found. We need an efficient algorithm to solve this problem.

3.2. The Rotation Invariance Algorithm

For two corresponding points on one shape and its rotation, we can find that one shape context vector can be derived by circularly shifting another shape context vector. The shifting position is a multiple of the number of bins for *log r* in the log-polar space. By using this kind of shifting on a shape context vector, there are different shape context vectors as many as the number of bins for θ. Each shape context vector corresponds to one rotation that can be realized with respect to the original shape in the log-polar space.

From the discussion above, we consider that all shape context vectors that can be derived by circular shift represents the corresponding points for one shape and all of its rotations. Therefore, these shape context vectors should be clustered together and replaced with the same shape label. For this purpose, we propose the rotation invariance algorithm as follows.

First, the *absolute distance* and *representative shape context vector* for 2 shape context vector in our log-polar space:

For two shape context vectors $X = (x_1, x_2, \ldots, x_{60})$ and $Y = (y_1, y_2, \ldots, y_{60})$, the absolute distance of X and Y is denoted as D_{XY} and its computation is

$$D_{XY} = \min_i \|X - R_i\| \quad , \quad 1 \le i \le 12 \text{, where } \begin{array}{l} R_1 = Y = (y_1, y_2, y_3, y_4, y_5, y_6, \ldots, y_{60}) \\ R_2 = (y_6, y_7, y_8, \ldots, y_{60}, y_1, y_2, y_3, y_4, y_5) \\ R_3 = (y_{11}, y_{12}, y_{13}, \ldots, y_5, y_6, y_7, y_8, y_9, y_{10}) \\ \vdots \\ R_{12} = (y_{56}, y_{57}, y_{58}, y_{59}, y_{60}, \ldots, y_{53}, y_{54}, y_{55}) \end{array}$$

The representative shape context vector is denoted as Y' and defined as

$$Y' = \arg\min_{R_i} \|X - R_i\| \quad , \quad 1 \le i \le 12$$

We use the absolute distance of two shape context vectors to measure the similarity of their shape context. The smaller the absolute distance, the more similar the two shape contexts are. We can see the absolute distance will be zero for two shape context vectors when their shape contexts are derived from corresponding points of a shape and its rotation. The steps for the rotation invariance algorithm are further described as follows.

Step 1: When there are N shape contexts in the database, the user decide the number of clusters, k.

Step 2: Select k shape contexts arbitrarily, and consider them as the centroid of each cluster.

Step 3: Use our proposed method to calculate the distance between each shape context and the centroid. That is, fixing the vector of centroid and rotating the vector of shape context by a multiple of a constant. Compute the minimum distance between rotated vectors and centroid, and record the rotated vector which is minimal.

Step 4: Group shape contexts according to the minimal distance. Compute the shape context which is closest to the centroid and include it into the group.

Step 5: If a shape context switches its cluster in step 4, update the centroid of the cluster:

$$C_i = \overline{\sum_{Y' \in \text{group } i} Y'} \quad , \quad 1 \le i \le k$$

Step 6: Repeat step 3, step 4 and step 5 until all of them are optimally clustered.

4. Experimental Results

Up to now, we are incapable of proving theoretically that there is a shape 3D model retrieval system is optimal. The most comprehensive comparative study of retrieval algorithms for 3D models is based on the now well-known Princeton Shape Benchmark [14]. The test set of the Princeton Shape Benchmark database contains 907 3D models in 92 categories. We retrieve the most similar 3D models for each query 3D model from the set by using the proposed 3D retrieval framework.

The average precision versus recall [15] is used to measure the retrieval performance in this experiment. We have compared the existing visual similarity system [16] with rotation invariance and without rotation 3D models in the proposed system as shown in Figure 2. The experimental results have shown that our system is effective and has better retrieval results than the existing systems.

5. Conclusions

In this paper, a 3D model retrieval system based on viewpoint representation is proposed. In this system, the view based approach is robust against rotation. The proposed rotational-invariant algorithm is adopted in the vector quantization process of shape contexts for the image database and groups similar shape contexts of images in the same cluster whether images are rotated or not. Therefore, the similar shapes will have some shape labels to the query object and can be retrieved by our system. This means our system has better retrieval performance than the other existing methods. The experiments on Princeton Shape Benchmark set confirm that our system has higher precision at all recall levels.

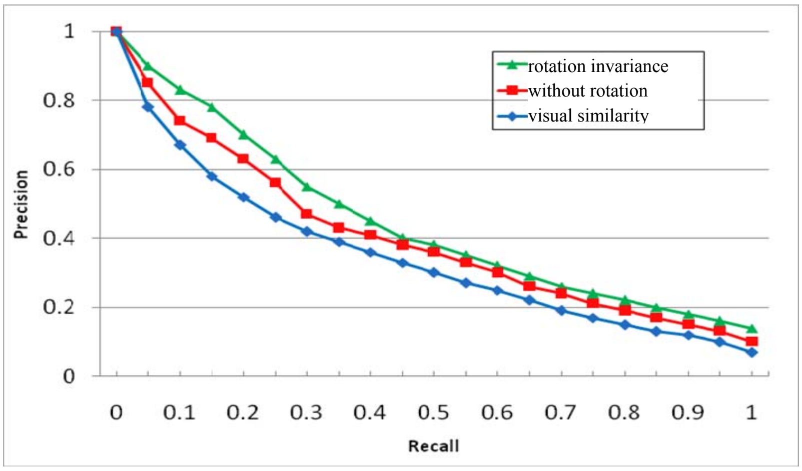

Fig. 2: Average precision vs. recall results from query with all categories.

6. References

[1] W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," *ACM Multimedia System*, vol. 7, pp. 184–198, 1999.

[2] F. Jing, M. Li, H. J. Zhang, and B. Zhang, "Relevance Feedback in Region-Based Image Retrieval," *IEEE Transactions Circuits and Systems for Video Technology*, vol. 14, pp. 672 – 681, 2004.

[3] F. Jing, M. Li, H. J. Zhang, and B. Zhang, "An Efficient and Effective Region-Based Image Retrieval Framework," *IEEE Transactions* Image *Processing*, vol. 13, pp. 699 – 709, 2004.

[4] J. Li, J. Z. Wang, and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval," *Proceedings of the 8th ACM* International *Conference on Multimedia*, pp. 147 – 156, 2000.

[5] S. Belongie, C. Carson, H. Greenspan, and J. Malik, "Recognition of images in large databases using color and texture," IEEE *Transactions Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026–1038, 2002.

[6] I. Biederman, "Recognition-by-components: a theory of human image understanding," *Psychological Review*, vol. 94 no. 2, pp. 115-147, 1987.

[7] P. Shilane, P. Min, M.K., Funkhouser, T, "The Princeton Shape Benchmark," *Proc. of International Conference on Shape Modeling and Applications*, pp. 167-178, 2004.

[8] Tangelder, T., Veltkamp, R, "A Survey of Content Based 3D Shape Retrieval Methods," *International Proc. of Shape Modeling*, pp. 145–156, 2004.

[9] D.-Y. Chen, X.-P. Tian, Y.T.S., Ouhyoung, M., "On Visual Similarity Based 3D Model Retrieval," *Computer Graphics Forum*, pp. 223-132, 2003.

[10] Vranic, D., "An Improvement of Rotation Invariant 3D Shape Descriptor Based on Functions on Concentric Spheres," *Proceedings of the International Conference on Image Processing*, pp. 757-760, 2003.

[11] Shum, H., "On 3D Shape Similarity," *Proc. of Computer Vision and Pattern Recognition*, pp. 526-531, 1996.

[12] Kang, S., Ikeuchi, K., "Determining 3D Object Pose Using the Complex Extended Gaussian Image," *Proc. of* Computer *Vision and Pattern Recognition*, pp. 580-585, 1991.

[13] T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Y. Wu, "An Efficient *k*-means Clustering Algorithm: Analysis and Implementation," *IEEE Transactions Pattern Analysis and Machine Intelligence*, pp. 881-892, July 2002.

[14] P. Shilane, P. Min, M. Kazhdan, and T. Fucnhouser, "The Princeton Shape Benchmark," *Proc. Shape Modeling* International, pp. 167 – 178, 2004.

[15] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Reading, MA: Addison-Wesley, 1999.

[16] Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen, and Ming Ouhyoung, "On Visual Similarity Based 3D Model Retrieval," *Computer Graphics Forum*, vol. 22, Issue 3, pp. 223–232, 2003.