3D Object Retrieval by Shape Similarity

Jeong-Jun Song and Forouzan Golshani

Department of Computer Science and Engineering,
Arizona State University,
Tempe, AZ, 85287-5406
{jjsong, golshani}@asu.edu
http://media.eas.asu.edu

Abstract. We introduce a method for shape similarity based retrieval in 3D object model database. The proposed method leads us to achieve effectiveness and robustness in similar 3D object search supporting both query by 3D model and query by 2D image. Our feature extraction mechanism is based on observation of human behavior in recognizing objects. Our process of extracting spatial arrangement of a 3D object by surface point distribution can be considered as using human tactile sensation without visual information. On the other hand, the process of extracting 2D features from multiple views can be considered as examining an object by moving viewpoints(camera positions). We propose shape signatures for 3D object model by measuring features of surface point and the shape distance distribution from multiple views of 3D model. Our method can be directly applied to industrial part retrieval and inspection system where different geometric representations are used.

1 Introduction

One of the most important points in designing interactive 3D multimedia systems is how to handle 3D data in an efficient manner. A number of research groups in the field of computer graphics and computer vision have attempted to find efficient ways of representing 3D objects, specifically, to identify each object based on its geometric properties. There are two main three-dimensional object representation schemes. One is called "object-centered representation", and the other is "viewer-centered representation scheme". Object-centered representation often uses explicit three-dimensional descriptions of objects of interest in Euclidian space like solid geometry and surface-based representation. On the other hand, viewer-centered representation usually relies on a collection of range or intensity images of an object as the implicit description of its shape. Since features extracted from a 2D image do not immediately correspond to the object models, 3D to 2D or 2D to 3D transformations must be made before observed features can be matched with the 3D object model. The multi-view models can be quickly searched for matching features but require large amount of storage to keep numerous 2D images from multiple viewpoints. How to derive the appearance of an object from a novel viewpoint without keeping too many views is the main difficulty. Since we are interested in the design of a multimedia retrieval

R. Cicchetti et al. (Eds.): DEXA 2002, LNCS 2453, pp. 851–860, 2002.

system that provides various query methods for efficiency and robustness, the speed of retrieving object model in the database might be more important issue than storage issue. We follow the 3D object-centered representation scheme as a main object representation but also utilize 2D views from each sampled viewpoint to extract additional useful features that may enhance shape identification. Once features from projection images have been extracted, it is not necessary to keep them for later use. To obtain boundary information from 2D views of 3D object we adapt one of the 2D shape descriptors, Fourier Descriptor (FD), that have been used in many image-processing applications. Moment invariance and Curvature Scale Space are also well known shape descriptors currently available. Our object-centered feature extraction method, called histogram representation of distance distribution, is a kind of spherical depth map with a surface approximation.

This procedure involves sampling surface points from the object and measuring the depth between surface intersection points and their corresponding points on the unit sphere. For the object composed of many triangles with small variation of triangle size and shape, it may not be necessary to resample the surface points. We restrict our experiment of resampling to objects having small number of triangles and large variation in triangle size. Histogram representation of this spherical depth map can be a characteristic of spatial arrangement of the object of interest in 3D co-ordination system. Next, we explore shape distance distribution based on observation of Human Visual System (HVS) mechanism. HVS is well studied in cognition and psychophysical sciences. Examples include Ullman's High-level Vision [1] and Schiffman's Sensation and Perception [2]. From selected viewpoints, we extract shape features and compute Euclidian distance from them. These shape signatures are compared to one another during query processing to find similar object in the 3D object model database. The importance of this approach becomes obvious when we reflect on HVS. Consider how humans normally study three-dimensional object. We study the object by rotating the object and examining different views of the object. That is, we gather characteristic information while changing the viewpoints and infer the object appearance without explicitly constructing 3D alignment.

The organization of this paper is as follows. First, we will discuss background and related work on shape representation and matching in section 2. Next in section 3, we will introduce shape feature extraction methods based on object-centered methods and multiple views. Section 4 presents experimental result. In section 5 we discuss conclusions.

2 Related Work

There has been a great deal of study for shape descriptors in 2D images, ranging from simple geometric attributes to various transformation techniques. These techniques can be categorized as reconstructive and non-reconstructive techniques, depending on whether or not they allow the original shape of the object to be approximated. A good categorization of 2D shape description technique can be found in [3][4]. The choice of specific description techniques depends totally on the nature of applications. Shape similarity has also been studied for 3D geometric model of the objects in computer vision and other object recognition areas. In most cases, they utilize range data to compare 3D local surface information with existing 3D object model stored in the database. The most widely used 3D object model description in computer graphics is based on geometrical information such as vertex points, surface point cloud, edge lists, facets and normal vector information. 3D modelling software such as CAD applications utilizes a solid modelling concept along with surface description. To be useful in large databases especially content-based retrieval system, feature extraction techniques should be simple, inexpensive and robust. Surface points distribution has some simple and efficient characteristics. Since Gauss introduced a representation method that maps normal vectors of surface patch to the sphere [5], related representation method have been developed as extensions of this mapping scheme. Extended Gaussian Image (EGI) [6] maps mass to the corresponding point on the Gaussian sphere. This mass distribution depends on the shape of the object. Distributed EGI (DEGI) [7] has been derived to avoid the problem for non-convex object of EGI. This method tessellates viewing sphere and recalculates a partial EGI for each viewing direction to determine the attitude of a non-convex object.

The EGI uniquely defines a convex polyhedron. However, non-convex object that has more than two separated region with the same surface normal may have the same representation with any convex object. A direct way of expanding 2D descriptor to 3D is to consider the 3D shape as composition of infinite 2D slice stack along a given direction. This is a common technique in medical applications. The other straightforward approach is to describe 3D object with multiple projection images. Aircraft detection is one of the applications utilizing this approach. Other 2D and 3D shape descriptors that are currently used in different applications will be discussed in the next section. Distributed EGI, a partial EGI for each viewing direction is considered to avoid this drawback. The hidden area, which is occluded, does not contribute EGI mass on Gaussian sphere.

Complex EGI (CEGI) [8] is introduced which measures the distance from the arbitrary origin to each surface patch to discriminate non-convex object having the same EGI representation. The weight at each point is complex number, Ae^{jd} whose magnitude is the surface area and whose phase is the distance over the Gaussian sphere. The complex weight associate with a surface patch is, where A is the area of a patch with surface normal, the normal distance d to a fixed origin. Since Complex EGI has no surface location information more than two parts on an object may be mapped on the same point of the sphere with the same weight.

Spherical Attribute Image (SAI) [9] provides a method to make one to one mapping between a non-convex object surface and a spherical surface. Most of 3D models are represented by unstructured triangular patches called free-form surface. SAI method deforms tessellated sphere called geodesic dome to original object surface as close as possible before extracting Gaussian curvature from it. This is a process that provides structured mesh without surface segmentation. The distribution based on the simplex angle is referred to as the Spherical Attribute Image.

3 Shape Features

Our approach mimics the behavior of human object recognition by providing features from 2D visual information and 3D spatial arrangement of objects. Humans use tactile and visual sensations to recognize an object of interest. Using three-dimensional spatial arrangement from distribution of surface points can be considered as touching object by hands to recognize its global shape whereas using visual information such as 2D contour information from different viewpoints can be considered as exploring object by rotating object to recognize its appearance. The following figure in Fig. 1 shows overview of our proposed retrieval process based on spatial features. The concept of aligning a 3D model with its principal axes may play an important role in 3D object feature analysis. Since the bounding box of the 3D object may vary by its initial object pose, the object alignment by its principal axes eliminate rotation normalization process in 3D object feature analysis. The fundamental problem to obtain principal axes of the 3D object lies on its computational complexity. To improve this problem, some methods have been proposed such as 1) Principle Component Analysis(PCA) with vertex weight [10], 2) PCA with weight proportional to triangle area [11]. Another problem is that the different size of triangles consisting of the object surface may cause widely varying normalized coordinate frame for

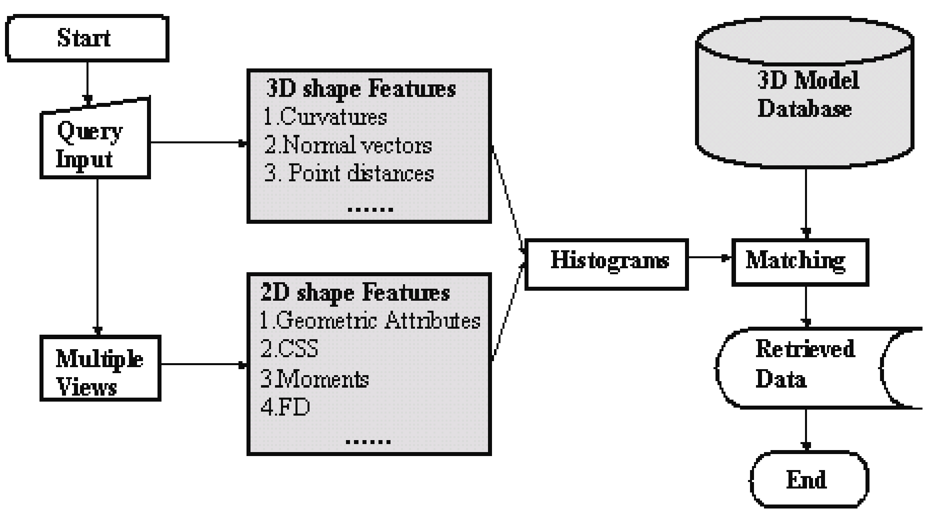

Fig. 1. Overview of proposed retrieval process. CSS: Curvature Scale Space, FD: Fourier Descriptor

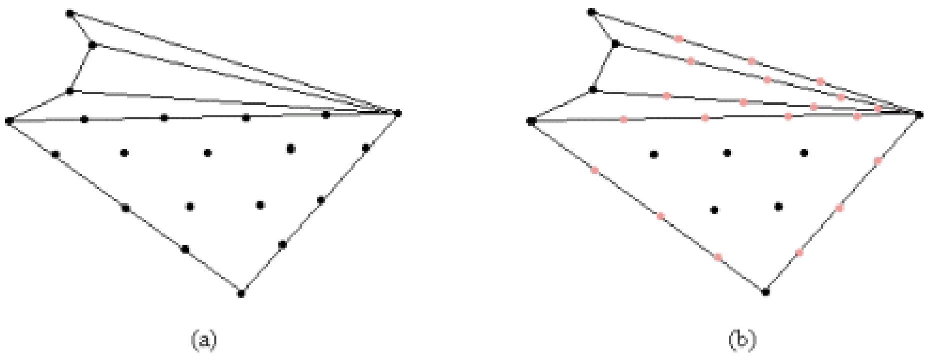

Fig. 2. (a) Problem of unbiased point distribution caused by various shape of triangle, (b) Biased point distribution after adding points

models that are identical. A typical method to generate unbiased random points with respect to the surface area of a polygon model is to subdivide triangles and obtain inner points. However, the various shape of triangle cause the biased point distribution. Our method improves it by adding points on triangle edges as well as faces as shown in Fig. 2. We first obtain inner points by applying recursive subdivision algorithm and add points on triangle edges recursively.

3.1 Depth Distribution

In order to make consistent environment for feature extraction we first generate unit sphere, where radius equals 1 (r = 1) and locate each object model in the center of the unit sphere. To make every object fit into unit sphere we need to find center of each object model and normalize it for size. The sampling points on a unit sphere can simply be obtain by subdividing polyhedron such as octagon, dodecahedron and icosahedrons. Once we determine the number of sampling points and find coordinates of each sampling point on the surface of the unit sphere, we shoot imaginary ray from each sampling point to the center of unit sphere. The length of each ray equals to 1, the radius of sphere. Since we locate 3D object model at the center of unit sphere, each ray is certain to intersect with the object. We compute intersection points and calculate distance between the starting point of the ray and the intersection point on the object surface. We use the heart of ray tracing technique to determine the intersection point of a ray with an object. To accomplish this task we use the parametric representation of a vector. Each point along the ray from A(x1, y1, z1) to B(x2, y2, z2) is defined by some value t such that

$$\left\{ \begin{array}{l} x = x_1 + (x_2 - x_1)t = x_1 + it \\ y = y_1 + (y_2 - y_1)t = y_1 + jt \\ z = z_1 + (z_2 - z_1)t = z_1 + kt \end{array} \right\} \tag{1}$$

Since we are using triangle mesh format for object models, we need to consider ray/plane intersection. A plane is a geometric entity that divides the space through which it passes in two. A plane, unlike a sphere, continues on infinitely; that is, it is unbounded. We begin with the equation of a plane, which is defined as:

$$A(x_1 + it) + B(y_1 + jt) + C(z_1 + kt) + D = 0 \tag{2}$$

Then the intersection is given by:

$$t = \frac{-(Ax_1 + By_1 + Cz_1 + D)}{(Ai + Bj + Ck)} \tag{3}$$

The point of intersection is calculated by substituting the value of t back into the ray equation. The next step is to determine if the intersection point is inside the polygon (triangle in our case). A number of different methods are available to solve this problem. One such algorithm works by shooting a ray in an arbitrary direction from the intersection point and counting the number of line segments crossed. If the number of crossings is odd, the point is inside the polygon; else it is outside. This is known as Jordan curve theorem. The detail explanation of ray tracing technique and Jordan curve theorem can be found in [12]. Now, we can calculate each distance between ray starting point and intersection point on the object surface and display each distance in a graph. This process corresponds to deformation of sphere to object surface as used in SAI. The advantage of our method is that we can get points from hidden surface. The fast ray-tracing algorithm can reduce computation time. To represent this distance distribution as a histogram we may set bin size to 0.1 for the distance (min 0.0, max 1.0) so that each distance belongs to one out of ten bins.

3.2 Spatial Arrangement Estimation

In this section, the histogram representations of 3D object features we devised for shape discrimination are described. The probability distributions of 3D object using distances of surface point pairs are discussed in [13]. In that paper, the authors utilized simple geometric attributes of surface points such as distance of two random points, angle between three random points, area of triangle between three random points and volume of tetrahedron between four random points and called them shape functions. We propose the way to estimate spatial arrangement of the object shape by utilizing Discrete Curvature and Normal Vectors as well as geometric attributes. The first step is to compute the Discrete Curvature of vertex points of triangle mesh. We use the Discrete Gaussian Curvature k which is defined as following Formula.

$$k = \frac{c}{A} \tag{4}$$

where c is the complete angle and A is barycentric area of triangles containing the corresponding vertex. The complete angle can be define as

$$c = 2\pi \sum_{i=1}^{m} \theta_i \tag{5}$$

and the barycentric area can be obtained from

$$A = \frac{1}{3}\sum_{i=1}^{m} t_i \tag{6}$$

Where m is number of neighborhood triangles containing the corresponding vertex, θ_i is adjacent angles between triangle edges containing the vertex and t_i is area of i^{th} triangle. The curvature value paly an important role in sorting the surface points in an order. Simply, we can measure distances from random pairs of surface points, differences of Discrete Curvature values and angles of normal vectors. Then, we produce three histograms based on the measurements. These histograms are turned out to be three different signatures of the same 3D object. In order to reduce computation time to generate signatures we select some surface points which have high curvature values. Then, the computations for generating signatures can be performed only for those selected surface points. In later case every pairs of selected points can be involved to generate shape signatures.

3.3 Shape Signature from Multiple Views

According to the recent psychophysical finding human perceives three-dimensional object from multiple view of two-dimensional shapes not from construction of 3D alignment. Consider how we study three-dimensional object in our real life. We study an object by rotating it, or changing the viewing angles. Shimon Ullman argued for the theory that view-independent object recognition by human is not based on an internally stored model of the object, but upon the outputs of neurons that respond to various two dimensional views of that object. In this section we propose a shape signature of 3D object models from multiple views for shape discrimination. We compute circularity of each 2D shape from different viewpoints and measure the shape distance between them. The circularity can be obtained from the ratio of boundary and area. A simple geometric shape like circle, rectangle can be used as the basis of various shapes. Then, the distances of simple geometric attributes between the basis and each shape from different viewpoints can be measured. This distance histogram generated from the distance valuses is considered as shape signature of the 3D model. Alternatively we can also measure the 2D shape distances from a pair of randomly selected viewpoints and demonstrate the probability distribution of them for an object signature. Many 2D shape features can be used for our approach. Fourier Descriptor(FD) allows us to compactly describe 2D shape from arbitrary viewpoints. Curvature Scale Space and Moment invariants can be used as the other 2D shape descriptor as well.

4 The Experiments

In this experiment, we first generated 2D binary projection images of Boeing-747 aircraft from different viewing directions as shown in Fig. 3. We utilized basic

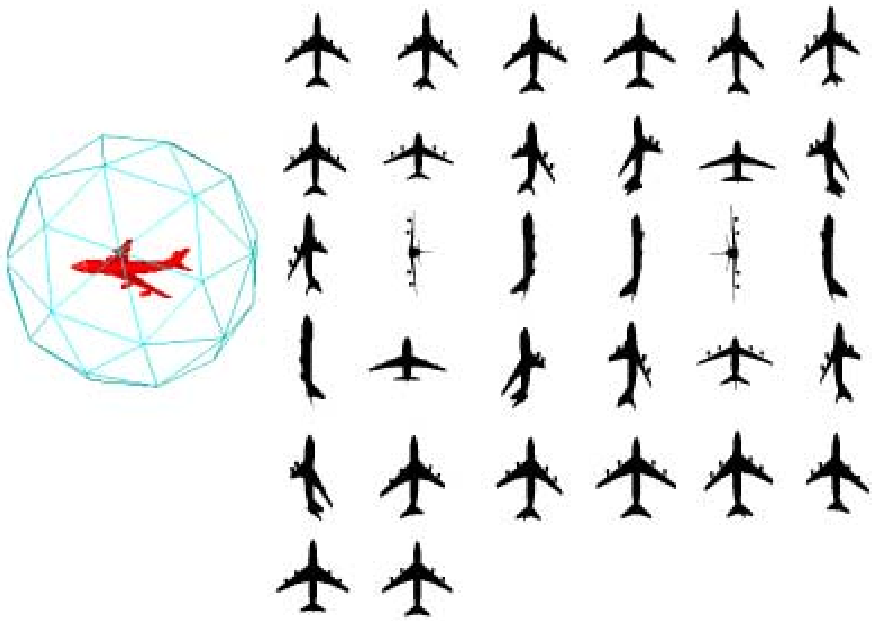

Fig. 3. Boeing 747 aircraft and its viewing sphere on the left and 2D projection binary images from different viewing directions on the left.

geometric attributes. For more reliable feature we will apply Curvature Scale Space and Moment Invariant features in future experiment. If we assume that 3D object have been initially located by principle axis we can directly consider the graph as a signature of the object model. However, since we have eliminated a procedure to find principle axis for the 3D model, the shape distance histogram can be used as the signature instead. Fig. 4(a) shows graph of geometric attribute (circularity) taken from 32 different viewpoints for Boeing 747 aircraft object. After computing circularity from each 2D images we can obtain distance from each pairs of circularity. This produces 1024 shape distances respectively. Then, histogram of shape distances for each geometric attribute will be obtained from them, which has 10 bins as shown in Fig. 4(b). Alternatively, shape distances can be obtained from random pairs of views.

5 Discussion and Conclusion

1D representation of the boundary of 2D shape has been developed in the context of image analysis for several decades. Our object-centered feature extraction approach can be conceptually extended from 2D shape feature extraction such as function of tangent angle and arc length, distance between shape centroid and boundary points, complex function using arc length parameter value. The dominant characteristic of our approach compared with others [14][15][16][17]

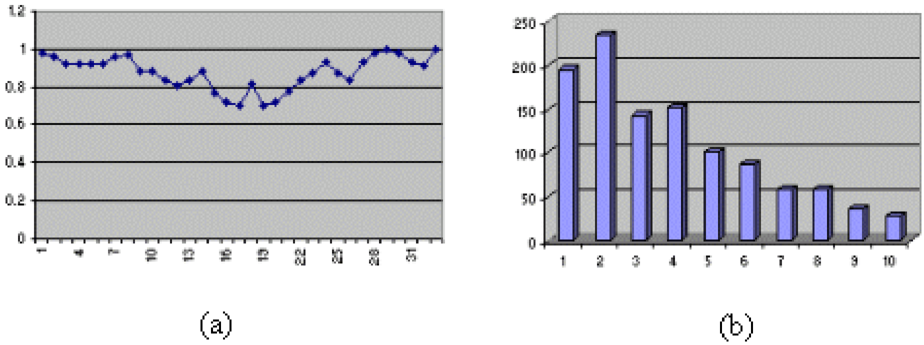

Fig. 4. (a)Graphs of circularity for Boeing 747 model x-axis: 32 different views y-axis: circularity. (b)Histogram of shape distance by circularity.

can be found in providing both object-centered and viewer-centered features by observation of human behavior for the object recognition. Especially, our view-centered approach mimics HVS to characterize 3D object by drawing several representative 2D view as appearance of a 3D object. Most view-based approach restrict its views to one viewing plane, namely, those generated be "walking around" the object without considering object pose which differ from the initial scanning. They usually assume that the object poses in stable condition, which is very ambiguous for object orientation. The main drawback is that it does not provide correspondences for arbitrary views of the object. In this paper we have proposed shape descriptors for general 3D object models based on object-centered and view-centered representation scheme respectively. We introduced 1D representation of the 3D shape using features of surface points. As another approach based on observation of HVS we explained a way to measure 2D shape distance from a pair of viewpoints and how to construct a histogram. These shape description method can be incorporated with the 3D shape distance measure that has geometrical differences in 3D model database. To be able to provide 2D image query interface our system need to develop efficient grouping algorithm based on 2D view similarity to reduce number of views of an object and clustering algorithm to hand large model database.

References

1. Shimon Ullman: High-level Vision, Object Recognition and Visual Cognition. The MIT Press(1996)
2. Harvey R. Schiffman: Sensation and Perception, An Integrated Approach. John wiley and sons, Inc.(1995)
3. Babu M. Mehtre, Mohan S. Kankanhalli, Wing F. Lee: Shape Measures for content based image retrieval: A comparision. Information processing and Management, Vol. 33, No.3 (1997) 319-337
4. Sven Loncaric: A survey of shape analysis Techniques. Pattern Recognition, Vol.31, No.8(1998) 983-1001

5. Gauss, K.F.: General Investigations of Curved Surfaces. Raven Press, New York(1965)
6. Horn, B.K.P.: Extended Gaussian Image. Proc. of IEEE, Vol. 72, No.12 December(1984)1671-1686
7. Ikeuchi, K. and Horn, B.K.P.: Picking up an Object from a Pile of Objects. in Robotics Research, The First International Symposium, J. M. Brady and R. Paul, The MIT Press, Cambridge, Massachusetts(1984) 139-162
8. Kang, S.B. and Ikeuchi, K.:The Complex EGI: New Representation for 3-D Pose Determination, IEEE Trans Pattern Analysis and Machine Intelligence, Vol. 15, No.7, July (1993) 707-721
9. Dellingette, H., Hebert, M., and Ikeuchi, K., " A Spherical Representation for the Recognition of Curved Objects," Proc. Intern. Conf on Computer Vision, May (1993) 103-112
10. D.V. Vranic and D. Saupe: Tools for 3D-object retrieval: Karhunen-Loeve Transform and spherical harmonics. submitted to IEEE Workshop on Multimedia Signal Processing, Cannes, France, Oct.(2001)
11. Paquet, E., Rioux, M.: A content-based search engine for VRML databases. Computer vision and Pattern Recognition Proceedings. IEEE Computer society conference(1998) 541-546
12. Andrew S. Glassner: An Introduction to Ray Tracing. ACADEMIC PRESS (1989)
13. R Osada, T Funkhouser, B Chazelle, and D Dobkin, "Matching 3D Models with Shape Distributions" ShapeModeling International , Genova, Italy, May, (2001)
14. C.C. Chang, S.M. Hwang and D.J. Buehrer: A shape recognition scheme based on relative distances of feature points from the centroid. Pattern Recognition 24(1991) 1053-1063
15. C. Zhang and T. Chen: Indexing and Retrieval of 3D models Aided by Active Learning. Demo on ACM Multimedia, Ottawa, Ontario, Canada, Oct.(2001) 1-5
16. C. M. Cyr, A. F. Kamal, T. B. Sebastian, and B. B. Kimia:2D-3D Registration based on Shape Matching. Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis(2000) 198-203
17. C. M. Cyr and B. B. Kimia: 3d object recognition using shape similarity-based aspect graph. In ICCV, pages Accepted – To appear, (2001)