3—33

RECOGNITION OF 3-D FREE-FORM OBJECTS USING DISTANCE-SUPPORTED SHAPE INDEX

Dong-O Kim[1], Sang Wook Lee[2], and Rae-Hong Park[1]

[1]Dept. of Electronic Engineering, [2]Dept. of Media Technology
Sogang University

ABSTRACT

The distance-supported shape index is proposed for the recognition of three-dimensional (3-D) free-form objects. It is constructed by incorporating extra global information into the locally computed shape index (SI) [4,5]. The extra information is the distance to the local region from: (1) the object's center of mass and (2) the major axis of object elongation. The major axis is defined as the eigenvector corresponding to the largest eigenvalue of the covariance matrix of 3-D object points. The cross entropy is used to measure the similarity between the histograms of distance-supported SIs. Experimental results from real range images show many instances of increased discriminating power of the distance-supported SI compared to the SI.

1. INTRODUCTION

Recognition of 3-D objects has been a topic of active research in computer vision, and various approaches to the representation of 3-D objects for recognition have been suggested [1]. Most of the representation schemes have adopted some form of surface/volumetric models, or local/global features such as surface normals, moments, spherical harmonics, 3-D curves, splashes and shape indices [2-9].

When rotation- and scale-invariant local shape features are computed in many parts of an object, the simplest way of utilizing the information would be to use only their distribution for recognition without establishing explicit geometric relations of the features within the object. It has been demonstrated that this is an efficient and effective way of representing arbitrary free-from objects [5]. It has also been shown that the use of local features alone is robust to partial occlusions [6]. Despite its efficiency and robustness to occlusions, however, the lack of geometric relations between the local features degrades its discriminating power substantially.

Full geometric matching, on the other hand, requires feature correspondences and is therefore computation intensive. It is not in general tolerant for partial occlusions. In this paper, we present an approach to incorporating partial geometric information into a local feature-based representation to improve its ability to discriminate objects. There have been efforts for encoding distance information from a designated origin to local features such as surface normals [8,9]. The novel aspect of our approach is that two different types of distance information are applied depending on the object elongation. One is the radial distance to a local region from a designated origin and the other is the axial distance from a designated axis. The object's center of mass serves as the reference origin. If an object is found to have a distinct major axis of elongation after a principal component analysis (PCA), our investigation indicates that the information based on the distance from the major axis is more effective in distinguishing between objects with similar local feature distributions.

The presented scheme is generally applicable to most of the local invariant features such as local moments, curvatures and spherical harmonics [7]. However, we use the shape index (SI) suggested in [4] since it is invariant to both rotation and scale and its effectiveness has been well examined [4-6].

The rest of the paper is organized as follows. Section 2 introduces two forms of distance-supported SIs and Section 3 discusses a strategy for employing representations based on the distance-based SIs and histogram matching. In Sections 4 and 5, experimental results and conclusions are presented, respectively.

2. DISTANCE SUPPORT FOR SHAPE INDEX

The SI is a quantitative measure of curvatured local shape and can be computed using principal curvatures. It was introduced by Koenderink and van Doorn [4] and modified by Dorai and Jain [5] as:

$$S_I(p) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\frac{\kappa_1(p)+\kappa_2(p)}{\kappa_1(p)-\kappa_2(p)}, \qquad (1)$$

where κ_1 and κ_2 denote the principal curvatures around the point p and $\kappa_1 \geq \kappa_2$. The SI as defined in Equation (1) has

[1, 2] C.P.O. Box 1142, Seoul, Korea 100-611 E-Mail: hyssop @sogang.ac.kr, slee@sogang.ac.kr, rhpark@ccs.sogang.ac.kr
This work was supported by grant number (R01-2002-000-00472-0) from the basic research program of Korea Science & Engineering Foundation (KOSEF).

the range [0,1], and every distinct surface shape corresponds to a unique value of SI. An exception is that planar surfaces are not defined in Equation (1) since both the principal curvatures are equal to zero. Hence, they are mapped to the value of 0.5 which happens to be the SI value of saddle shapes. The SI is invariant to rotation, translation and scale, and the 1-D histogram $H_S(S_I)$ computed with all the surface points, can be used for matching. We augment the shape representation by adding distance information from a designated center or axis in the object space.

It has been shown that encoding of distance information can increase the ability to disambiguate objects that have similar surface normal distributions [8,9]. We combine the radial distance from the object center d_r with the shape index S_I. The addition of this information results in a 2-D histogram $H_R(S_I, d_c)$. To maintain the scale invariance of S_I, the magnitude of the range data points are normalized with respect to the maximum radial distance. In other words, for the data points $\boldsymbol{x}_m=[x_m\ y_m\ z_m]^T$ $(1\le m\le M)$ the mean is computed as:

$$\boldsymbol{\mu} = \frac{1}{M}\sum_{m=1}^{M} \boldsymbol{x}_m, \tag{2}$$

and the data points are unbiased and normalized as:

$$\boldsymbol{\phi}_m = \frac{\boldsymbol{x}_m - \boldsymbol{\mu}}{\max_m \left\| \boldsymbol{x}_m - \boldsymbol{\mu} \right\|}, \tag{3}$$

when the radial feature distances are computed.

When an object is substantially elongated, the principal axis of elongation is stably defined from object points and we use the axial distance from the principal axis as another source of geometric information for recognition. The principal axis can be computed from PCA with the covariance matrix:

$$\mathbf{C} = \mathbf{A}\mathbf{A}^T, \tag{4}$$

where $\mathbf{A}=[\boldsymbol{\varphi}_1\ \boldsymbol{\varphi}_2\ \ldots \boldsymbol{\varphi}_m]$. From $\mathbf{C}\boldsymbol{e} = \lambda\boldsymbol{e}$, three eigenvalues λ_1, λ_2 and λ_3 ($\lambda_1\ge\lambda_2\ge\lambda_3$) and three eigenvectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are obtained, and the eigenvector $\boldsymbol{e}_1$ corresponding to λ_1 is the major axis of object elongation and thus taken as the reference axis. The addition of the information from the axial distance from $\boldsymbol{e}_1$, results in another 2-D histogram $H_A(S_I, d_a)$. Since the data points are normalized and the major axis captures the object orientation, both $H_R(S_I, d_r)$ and $H_A(S_I, d_a)$ are rotation and scale invariant. Fig. 1 illustrates the distances d_r and d_a in the 3-D object space.

We can possibly use more than one axial distance based on the three axes computed by PCA. In principle, the three axial distances from $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ would completely define a 3-D object geometrically. However, although there is a large class of objects with the dominant major axis, the second and the third axes are indistinctly defined for most of the common objects and the practical utility in using more than one axial distance is quite limited.

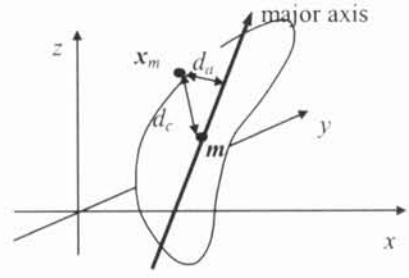


Fig. 1. Distances from object center and major axis.

3. OBJECT REPRESENTATION AND MATCHING

A 3-D object can be represented by one of the two 2-D distance-supported SI histograms $H_R(S_I, d_r)$ and $H_A(S_I, d_a)$ since the 2-D histograms contain the information about the 1-D SI histogram $H_S(S_I)$. For an elongated object, both H_R and H_A can be used for matching. However, our experimental studies presented in Section 4 suggest that H_A is generally more effective than H_R for elongated objects, and therefore only H_A can be reserved when $r_\lambda = \lambda_2/\lambda_1$ is sufficiently low. The extra parameter r_λ guides initial indexing by determining the type of the histograms for subsequent matching. The object pose can be determined in a similar fashion to that in [5].

When $H_R(S_I, d_c)$ and $H_A(S_I, d_a)$ are constructed with the same number of histogram bins as that for $H_S(S_I)$ to keep the computational load comparable, H_R and H_A carry only crude information about $H_S(S_I)$ due to the coarse quantization in the S_I direction. For instance, when the number of bins for H_S is 100, that for H_R should be 10×10 for the same size and computational cost for matching. Even in this case, our experimental studies show that H_R and H_A have more discriminating power than H_S in most cases, hence it is not necessary to keep H_S in an object's representation.

Histogram matching has been used for some object indexing and recognition approaches and a few matching methods have been suggested, such as histogram intersection, the χ^2 test and probabilistic matching [10]. In the work presented in this paper, we use the cross entropy as a measure of similarity [11].

The cross entropy $C(Q,P)$ of two 2-D histograms Q and P is defined as

$$C(Q,P) = \int\int Q(S_I,d)\log\frac{Q(S_I,d)}{P(S_I,d)}\,dS_I\,dd. \tag{5}$$

The symmetric cross entropy $(C(P,Q)+C(Q,P))/2$ is actually used for measuring similarity in our experiments.

4. EXPERIMENTAL RESULTS

Experiments are performed to investigate the effectiveness of the SI and the proposed distance-

supported representations. Figs. 2 and 3 show the range images (available at http://sampl.eng.ohio-state.edu/~sampl/data/3DDB/index. html) used in our experiments. The range images in Fig.2 are from the same object but with different pose and scale. The objects in Figs. 2 (b), (c), and (d) are rotated from the object in Fig. 2 (a) by 20 degrees about the x, y, and z axes, respectively. The range image in Fig. 2 (e) is scaled version of that in Fig. 2 (a). The objects shown in Fig. 3 are all different from that in Fig. 2 and from each other.

Fig. 4 (a) and (b) show the SI histograms from the range images in Figs. 2 and 3, respectively. The number of histogram bins is 100 for both Fig. 4 (a) and (b). All the SI histograms in Fig. 4 (a) are similar since they are from the same object and the SI is rotation and scale invariant. Since the objects shown in Fig. 3 mainly have ridge and dome surfaces while the object in Fig. 2 consists mostly of ridge and rut surfaces, all the SI histograms in Fig 4. (b) are substantially different from those in Fig. 4 (a). However, those in Fig. 4 (b) are quite similar to each other despite the fact that they are from different objects that look substantially dissimilar. Therefore, the SIs are not effective for distinguishing them. This is demonstrated in Table I where the cross entropy values between the SI histograms $H_S(S_I)$s for intra-distances (between the same object with different pose or scale) and those for inter-distances (different objects). As seen in Table I, the inter-distances between the objects in Fig. 3 are not much higher than the inter-distances between the same objects in Fig. 2. For higher discriminating power, it is desirable to have bigger differences between the intra- and inter-distances.

Table II (a) and (b) show the intra- and inter-distances but with the distance-supported SI histograms $H_R(S_I, d_r)$s and $H_A(S_I, d_a)$s. The total number of bins for these 11×11 2-D histograms is 121 and is comparable to that of $H_S(S_I)$. It can be seen that while the intra-distances increase marginally from $H_S(S_I)$s to $H_R(S_I, d_r)$s and $H_A(S_I, d_a)$s, the rates of increase for inter-distances are substantially higher for the inter-distances. This means that the different objects are much better distinguishable based on $H_R(S_I, d_r)$ and $H_A(S_I, d_a)$ than on $H_S(S_I)$. It should be noted that all the objects in Figs. 2 and 3 are highly elongated and the inter-distance cross entropy values from $H_A(S_I, d_a)$ are considerably larger than those from $H_R(S_I, d_r)$.

5. CONCLUSIONS AND FUTURE WORK

This paper presents a new method for incorporating partial information about global geometry into a local feature-based shape representation. The global information is provided in the form of local feature's radial distance from the object's center and axial distance from the major axis of object elongation. Our investigation suggests that the distance-supported shape index have a greater degree of discriminating power for recognition than shape index alone. It also shows that the axial distance-supported shape index is more effective than the radial distance-supported shape index for distinguishing elongated objects.

We are interested in investigating the effectiveness of the distance-supported shape index for recognizing objects with partial occlusions. Our preliminary study suggests that the distance-supported shape index is almost as tolerant for partial occlusions as the shape index.

REFERENCES

[1] P.J. Besl, "The free-form surface matching problem", *Machine Vision for Three-Dimensional Scenes*, H. Freeman, ed., pp. 25-71. Academic Press, 1990.

[2] R.J. Campbell and P.J. Flynn, "A survey of free-form object representation and recognition techniques," Computer Vision and Image Understanding, vol. 81, no. 2, pp. 166-210, Feb. 2001.

[3] F. Stein and G. Medioni, "Structural indexing: efficient 3-D object recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-14, no. 2, pp. 125-145, Feb. 1992.

[4] J. J. Koenderink and A. J. van Doorn, "Surface shape and curvature scales," *Image and Vision Computing*, vol. 10, no. 10, pp. 557-565, Oct. 1992.

[5] C. Dorai and A. K. Jain, "COSMOS- A representation scheme for 3D free-form objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-19, pp. 1115-1130, Oct. 1997.

[6] G. Hetzel, B. Leibe, P. Levi, and B. Schiele, "3D object Recognition from Range Images using local feature histograms", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR 2001*, vol. 2, pp. 394-399, Kauai Island, Hawaii, Dec. 2001.

[7] G. C. Sharp, S. W. Lee, and D. K. Wehe, "ICP registration using invariant features," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol PAMI-24, no 1, pp. 90-102, Jan. 2002.

[8] V. S. Nalwa, "Representing oriented piecewise C2 surfaces," in *Proc. IEEE International Conference on Computer Vision*, pp. 40-51, vol. 1, no. 2, Tarpon Springs, FL., Dec. 1988.

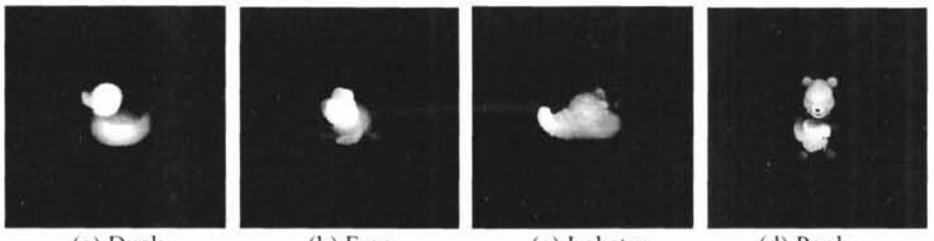
[9] S. B. Kang and K. Ikeuchi, "The complex EGI: A new representation for 3-D pose determination," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-15, no. 7, pp. 707-721, July 1993.

[10] B. Schiele and J. Crowley, "Representation without correspondence using multidimensional receptive field histograms," *Intern. Journal of Computer Vision*, vol. 36, no. 1, pp 31-52, Jan. 2000.

[11] D.-O. Kim and R.-H Park, "Classification of 3-D objects with curved surfaces based on the cross entropy between shape histograms," *Electronic Letters*, vol. 36, no. 19, pp. 1614-1615, Sept. 2000.

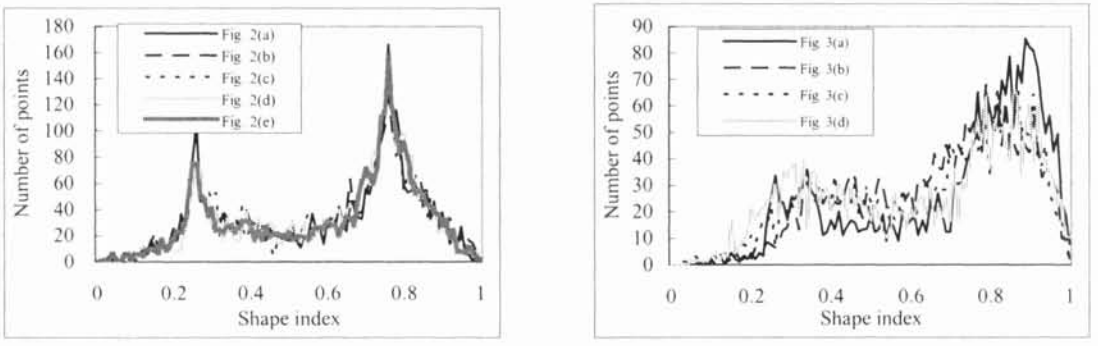
(a) Sculpture (b) θ_x=20° (c) θ_y=20° (d) θ_z=20° (e) Zoommed image

Fig. 2. Same images (original and transformed images)



(a) Duck (b) Frog (c) Lobster (d) Pooh

Fig. 3. Different images



(a) $H_S(S_I)$ histograms of the objects in Fig. 2 (b) $H_S(S_I)$ histograms of the objects in Fig. 3

Fig. 4. SI histograms of the objects in Figs. 2 and 3

Table I. Cross entropy between SI histograms ($\times 10^3$)

(a) Intra-distances

	Fig.2(a)	Fig.2(b)	Fig.2(c)	Fig. 2(d)	Fig.2(e)
Fig.2(a)	0	15.1	11.4	15.6	13.8
Fig.2(b)		0	12.2	13.1	16.0
Fig.2(c)			0	16.1	14.7
Fig. 2(d)				0	22.0
Fig.2(e)					0

(b) Inter-distances

	Fig.2(a)	Fig.3(a)	Fig.3(b)	Fig. 3(c)	Fig.3(d)
Fig.2(a)	0	114.8	60.7	81.7	67.8
Fig.3(a)		0	42.1	48.9	38.3
Fig.3(b)			0	12.5	19.2
Fig. 3(c)				0	38.5
Fig.3(d)					0

Table II. Cross entropy between the histograms $H_R(S_I, d_r)$s and $H_A(S_I, d_a)$s ($\times 10^3$)

(a) Intra-distances (between $H_R(S_I, d_r)$s / $H_A(S_I, d_a)$s)

	Fig.2(a)	Fig.2(b)	Fig.2(c)	Fig. 2(d)	Fig.2(e)
Fig.2(a)	0	15.2/21.9	10.0/19.7	8.5/28.1	25.7/26.3
Fig.2(b)		0	13.9/27.6	19.8/31.3	26.9/26.4
Fig.2(c)			0	16.6/22.3	29.7/24.5
Fig. 2(d)				0	23.2/29.0
Fig.2(e)					0

(b) Inter-distances (between $H_R(S_I, d_r)$s / $H_A(S_I, d_a)$s)

	Fig.2(a)	Fig.3(a)	Fig.3(b)	Fig. 3(c)	Fig.3(d)
Fig.2(a)	0	169.3/326.7	114.5/312.9	94.2/262.7	125.3/226.2
Fig.3(a)		0	81.8/90.4	115.0/196.3	48.5/157.2
Fig.3(b)			0	57.5/137.0	70.8/251.0
Fig. 3(c)				0	90.8/186.8
Fig.3(d)					0