화상에서의 각도 변화를 이용한 3차원 물체 인식

황 보 택 근†

요 약

컴퓨터 비전을 이용한 3차원 물체 인식은 카메라의 위치에 따라 화상에 투영되는 물체의 형상이 변하기 때문에 매우 복잡하고 어렵다. 따라서 컴퓨터 비전을 이용한 효과적인 인식 시스템을 구축하기 위해서는 각 3차원 물체에 있어서 유일하고 중요한 특성이 보는 각도에 따라 어떻게 변화하는가를 분석하고 이해하는 것이 매우 중요한 요소이다. 본 연구에서는 특징점들(landmarks)간에 이루어지는 각도 또는 3차원 다각형의 모서리(edge) 사이의 각도를 중요한 특성으로 선택하였고, orthographic 투영과 isotropic view orientation 아래에서 그러한 각도들의 보는 방향에 따른 화상에서의 변화를 2차원 결합 밀도 함수로 유도하였다. 본 논문에서 구한 수리적인 결합 밀도 함수는 통계적인 판단 규칙을 적용하여 효과적으로 물체 인식에 적용될 수 있다. 제안된 방법의 타당성 검토를 위하여 간단한 실험을 수행하였으며, 실험결과 본 방법이 매우 효과적인 것으로 나타났다.

View Variations and Recognition of 3-D Objects

Whangbo Taegkeun †

ABSTRACT

Recognition of 3D objects using computer vision is complicated by the fact that geometric features vary with view orientation. An important factor in designing recognition algorithms in such situations is understanding the variation of certain critical features. The features selected in this paper are the angles between landmarks in a scene. In a class of polyhedral objects the angles at certain vertices may form a distinct and characteristic alignment of faces. For many other classes of objects it may be possible to identify distinctive spactial arrangements of some readily identifiable landmarks.

In this paper given an isotropic view orientation and an orthographic projection the two dimensional joint density function of two angles in a scene is derived. Also the joint density of all defining angles of a polygon in an image is derived. The analytic expressions for the densities are useful in determining statistical decision rules to recognize surfaces and objects. Experiments to evaluate the usefulness of the proposed methods are reported. Results indicate that the method is useful and powerful.

1. Introduction

Most schemes for recognizing 3-D objects in two-dimensional images are model-based systems in which recognition requires matching features in the input image with sets of model features. In most situations the orientation of the object vis-à-vis the camera is not fixed, there is, therefore, considerable variation in the images of an object that are captured by a machine vision system. It is impractical to store all of

† 정 회 원 : 경원대학교 공과대학 전자계산학과
논문접수 : 1997년 5월 6일, 심사완료 : 1997년 10월 18일

these image variations and to attempt recognition by matching an objservation to these stored images. More feasible recognition schemes [1]-[3] store features sampled at different views and use these feature sets as model representations of the object. An understanding of feature variations with view orientation is fundamental to designing efficient recognition systems.

The importance of understanding the variation in geometric features is recognized. Burns *et al* [4] quantify the variations in angles, length, and relative position and plot extensive charts and histograms of these variations. Analytic expressions for variations are however not derived. Ben-Arie [5] develops probabilistic models for angles and distances but his functional expressions are mostly approximations. Binford et al [6] have also studied angle densities but have developed their *pdf* plots using numerical integration.

The ideas presented in this paper are continuation/ extension of previous and ongoing research in 3-D object recognition using view variation. The random orientation model (*rom*) was developed and applied to specific situations in [7][9]. In [10] the angles between the edges of the surface were selected as a set of significant characteristics of a polygonal flat surface and an analysis of the variations of the measured angles in a triangle was presented. The angles selected are not restricted to the polyhedral object, the spatial arrangement of landmarks (or easily recognized icons) on an object may constitute a unique characteristics of that object.

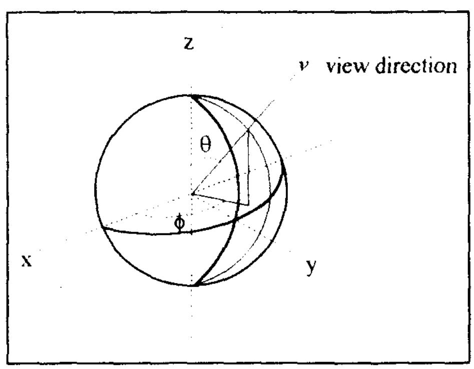
A significant characteristic of the decision rules using the variation of the measured angles is to distinguish amongst objects that appear in images, possibly with the same shape. However since such decision rule is statistical, it will not always give the correct result. A method to reduce the error probability is examining the more measured angles in an image. Some of new contributions of this paper are 1) the verification of the usefulness of the proposed decision rules through some experiments, 2) a mathematical framework and mathematical understanding of the view variation, and 3) the derivation of the joint densities of the defining angles of any polygon.

A brief description of random orientation model and angle measures in an image is presented in section 2. In section 3 the joint densities of the defining angles of polygon are derived. Optimum decision rules for recognition are outlined in section 4. Experiments to verify the utility of these rules are reported in section 5. Finally, in section 6 some conclusions are drawn.

2. Angle Measures in Images and Angle Density

The image captured is assumed an orthographic projection of the scene and that the image plane is normal to the view axis (Fig.1). Object views from different directions are equivalent to positioning the viewpoint, v, at different locations on the observation sphere. A location on the observation sphere, and thus a unique orientation of the object, is specified by the polar and azimuth angles, θ and ϕ, respectively.

In the random orientation model the position of the viewpoint on the observation sphere is randomly selected. This in turn implies that θ and/or ϕ are random. The probability density function (*pdf*) that describe the randomness in θ and ϕ is determined by the

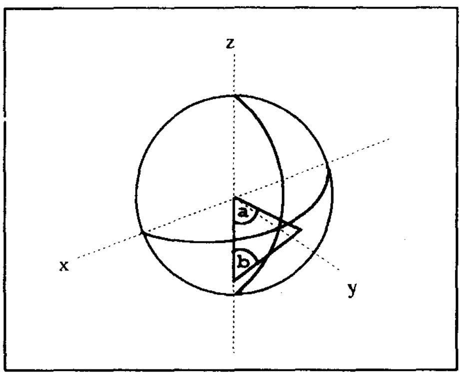


(Fig. 1) An observation sphere models camera and/or object orientation.

viewing situation and derived in [11] as

$$f_{\theta,\phi}(\theta,\phi)=\begin{cases}\frac{1}{4\pi}\sin\theta; 0\le\theta\le\pi,\ 0\le\phi\le 2\pi \\ 0 \quad ;\text{elsewhere}\end{cases} \tag{1}$$

A triangle is placed in the x-z plane at the center of an observation sphere (Fig. 2). The angles in the plane of this triangle are labeled, a, b and c. ⟨Table 1⟩ contains a list of the labeling conventions used.



(Fig. 2) A triangle within the observation sphere. The triangle is in the x-z plane.

⟨Table 1⟩ A catalog of the labels used in (Fig. 2).

Description	Convention	Example
Angles in the plane	Lower-case	a, b
Angles on the sphere (or in image)	Upper-case	A, B
Angles at the origin	Greek	θ, ϕ
Intersection points	Italics	*a*, *b*

Using the Jacobian method, the density function $f_{A,B}(A, B)$can be simplified as (see [10] or [11] for detail derivations)

$$f(A,B)=\frac{1}{2\pi\sin A\sin B\sin C}\left(\sqrt{\frac{1}{1-\left(\frac{2}{kp}\right)^2}}-1\right) \tag{2}$$

$;(0<A\le\pi,\ 0<B\le\pi-A,\ C=\pi-A-B)$

$$where\ k=\frac{\sin(A+B)\sin a\sin b}{\sin A\sin B\sin(a+b)}=\frac{\cot A+\cot B}{\cot a+\cot b},$$

$$p=\frac{1+l^2+k^2}{k^2}$$

3. Polygon Angle Densities

3.1 Triangle

The angle-pair density in equation (2) is accurate only if a correct point correspondence has been established between the image triangle and the model triangle. Suppose this image triangle is a projection of the model triangle at some unknown orientation. Without collateral information it is impossible to determine if the angle pair $(A\ B)$ in the image corresponds to (a b), or to (b c), or to (c a) in the model. It is reasonable to assume that there is only a 1/3 chance of selecting the correct correspondence. Therefore the angle-pair density for a given model T is

$$f(A,B/T)=\frac{1}{3}f(A,B/\text{a, b})+\frac{1}{3}f(A,B/\text{b, c})+\frac{1}{3}f(A,B/\text{c, a}) \tag{3}$$

where $f(A, B/\text{a, b})$ is the density in equation (2), and $f(A, B/\text{b, c})$ and $f(A, B/\text{c, a})$ are functional equivalent to the expcept that k and p are replaced with corresponding angles in the model.

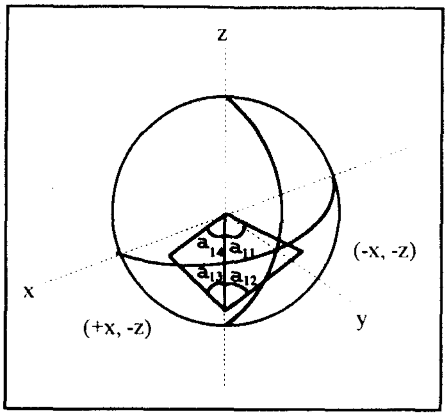
4.2 Quadrilaterals

A quadrilateral is just a combination of two triangles sharing one side, which is a diagonal of the quadrilateral. The diagonal of a quadrilateral is in negative z axis, and one triangle is in (+x, −z) plane and another triangle is in (−x, −z) plane as shown in (Fig. 3). Let a_{11}, a_{12} be the two angles of the triangle in (−x, −z) plane and a_{13}, a_{14} be the two angles of the triangle in (+x, −z) plane. The images of angles a_{11}, a_{12}, a_{13}, and a_{14} are denoted as A_{11}, A_{12}, A_{13}, and A_{14} respectively. The relationship between angle a_{13},

a_{14} and A_{13}, A_{14} is obtained analogously by using the geometry of spherical triangles.

$$\cot A_{13} = \csc\phi \sin\theta \cot a_{13} + \cot\phi \cos\theta \quad (4)$$

$$\cot A_{14} = \csc\phi \sin\theta \cot a_{14} - \cot\phi \cos\theta \quad (5)$$



(Fig. 3) A quadrilateral within the observation sphere.

The probability density of the angle vector $[A_{11}, A_{12}, A_{13}, A_{14}]$ given the correspondence ($a_{11} \rightarrow A_{11}$, $a_{12} \rightarrow A_{12}$, $a_{13} \rightarrow A_{13}$, $a_{14} \rightarrow A_{14}$) may be obtained using Bayes' rule:

$$f(A_{11} A_{12} A_{13} A_{14} \mid a_{11} a_{12} a_{13} a_{14}) = f(A_{13} A_{14} \mid A_{11} A_{12} ; a_{11} a_{12} a_{13} a_{14}) \cdot f(A_{11} A_{12} \mid a_{11} a_{12}) \quad (6)$$

Now the joint density of A_{11}, A_{12} is just equation (2) with the following substitutions:

$A_{11} = A$, $A_{12} = B$, $k_1 = k$, etc. ; i.e.

$$f(A_{11}, A_{12} \mid a_{11} a_{12}) = \frac{1}{2\pi \sin A_{11} \sin A_{12} \sin(A_{11} + A_{12})}$$

$$\left(\frac{1}{\sqrt{1 - \left(\frac{2}{k_1 p_1}\right)^2}} - 1 \right), \ where\ p_1 = \frac{1 + l_1^2 + k_1^2}{k_1^2}$$

Now A_{13} and A_{14} are completely determined by the other angles, hence the angle vector density may be written in terms of delta functions [12] as:

$$f(A_{11} A_{12} A_{13} A_{14} \mid a_{11} a_{12} a_{13} a_{14}) = \delta\left(A_{13} - \cot^{-1}(k_1 \cot a_{13} + l_1)\right) \cdot \delta\left(A_{14} - \cot^{-1}(k_1 \cot a_{14} + l_1)\right) \cdot f\left(A_{11} A_{12} \mid a_{11} a_{12}\right) \quad (7)$$

In a convex quadrilateral there is no way to correctly identify vertex with the (a_{11}, a_{14}) angles without additional information. An image vertex (A_{13}, A_{14}) may correspond to any of the model vertices (a_{11}, a_{12}), (a_{13}, a_{14}), (a_{21}, a_{22}), or (a_{23}, a_{24}), with only 14 chance of selecting the correct one. The *pdf* of a convex quadrilateral, Q, is therefore

$$f\left(A_{11} A_{12} A_{13} A_{14} \mid Q_{convex}\right) = \frac{1}{4} f\left(A_{11} A_{12} A_{13} A_{14} \mid a_{11} a_{12} a_{13} a_{14}\right) + \frac{1}{4} f\left(A_{11} A_{12} A_{13} A_{14} \mid a_{13} a_{14} a_{11} a_{12}\right) + \frac{1}{4} f\left(A_{11} A_{12} A_{13} A_{14} \mid a_{21} a_{22} a_{23} a_{24}\right) + \frac{1}{4} f\left(A_{11} A_{12} A_{13} A_{14} \mid a_{23} a_{24} a_{21} a_{22}\right)$$

where each term on the RHS is derived from equation (7).

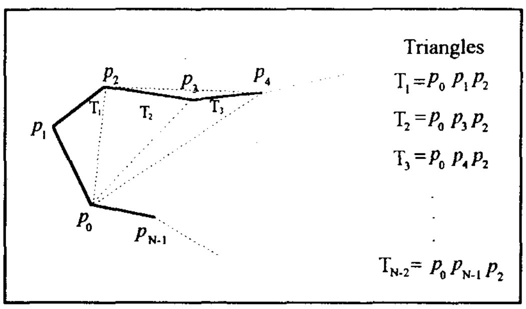
In a convave quadrilateral, there can only be one vertex in a quadrilateral with an inside angle greater than 180°. Also all views of this angle lie between 180° and 360°, i.e. a concave angle always remains a concave angle Therefore there is no ambiguity in identifying the concave vertex, and the quadrilateral density is

$$f\left(A_{11} A_{12} A_{13} A_{14} \mid Q_{convex}\right) = f\left(A_{11} A_{12} A_{13} A_{14} \mid a_{11} a_{12} a_{13} a_{14}\right)$$

3.3 N-gon

The *pdf* of a N-sided polygon can similarly be derived. The N-gon can be split into N-2 triangles by using the method illustrated in (Fig. 4), and therefore with 2(N-2) angles the N-gon is uniquely determined. There are however, N possible ways of dividing a N-gon into N-2 triangles. If an N-gon consists of only convex angles, then the angle vector density function is

$$f\left(A_{11},\ A_{12},\ldots,A_{1(2N-4)}\mid N-gon\ conex\right)$$
$$=\frac{1}{N}\sum_{i=1}^{N} f\left(A_{11},\ A_{12},\ldots,A_{1(2N-4)}\mid a_{i1},\ a_{i2},\ldots,a_{i(2N-4)}\right)$$
$$where\ f\left(A_{11},\ A_{12},\ldots,A_{1(2N-4)}\mid a_{i1},\ a_{i2},\ldots,a_{i(2N-4)}\right)=$$
$$f\left(A_{11}A_{12}\mid a_{i1},\ a_{i2}\right)\prod_{j=1}^{2N-4}\delta\left(A_{1j}-\cot^{-1}\left(k_i\cot a_{ij}+l_i\right)\right),$$
$$l_i=\frac{\cot A_{12}\cot a_{i1}-\cot A_{11}\cot a_{i2}}{\cot a_{i1}+\cot a_{i2}},$$
$$k_i=\frac{\sin(A_{11}+A_{12})\sin a_{i1}\sin a_{i2}}{\sin A_{11}\sin A_{12}\sin(a_{i1}+a_{i2})}$$



(Fig. 4) With the vertex labeled shown on an N-gon it is possible to define N-2 triangles as shown.

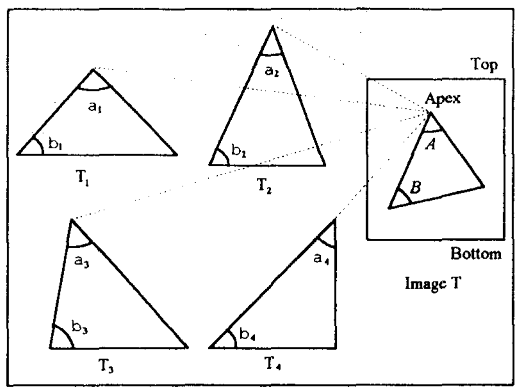
4. Decision Rule

The density functions that have been derived may be used to develop optimum decision rules for recognizing plane geometric shapes in images. In this paper exposition is restricted to the recognition of triangles; however the methods presented here may easily be extended to recognizing N-gons.

Decision rules depend on the a priori knowledge available. Two distinct situations exist in recognizing triangles in images. The more desirable situation is when point correspondence has been established between the points (or vertices) of an image triangle and the model triangles using connectivity or other information acquired by processing other parts of the image. The converse situation is when no correspondence has been established implying greater ambiguity.

4.1. Point Correspondence

Consider the situation where the triangles shown in (Fig. 5) represent, say, parts of the facade of various buildings and our objective is to process an image of one of these buildings and to recognize the shape of a triangle and thereby alos recognize the building.



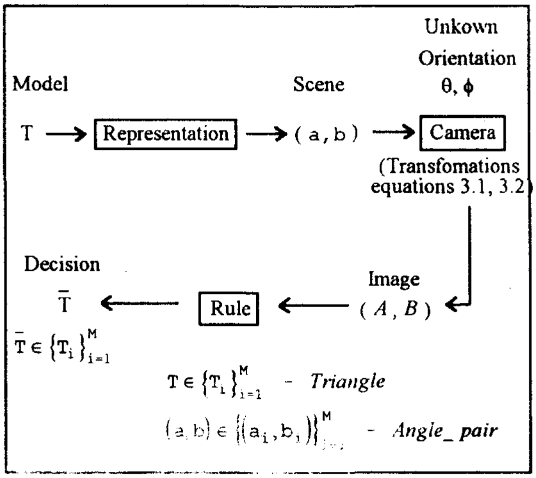
(Fig. 5) Model triangles T_1, T_2, T_3, and T_4 where the apex is recognizable in the image T.

A model of the decision problem is shown in (Fig. 6). The objective solution to this problem is the specification of a rule to convert the image angle-pair (A, B) measurements into a decision about which model triangle produced (A, B). An optimum solution is a decision rule which minimizes some reasonable cost function such as the probability of error, P(E).

The decision rule which minimizes the probability of error in the case when all model triangles (T_1, T_2,..., T_M) are equally likely is

$$\overline{T}=T_i\ if\ f\left(A,\ B\mid a_i,\ b_i\right)\geq f\left(A,\ B\mid a_j,\ b_j\right)\ \forall i\neq j \quad (8)$$

This rule may further be simplified by using equation



(Fig. 6) The decision problem model. T is one of the triangles from the set $\{T_1, T_2, \cdots, T_M\}$ and is characterized by the angle-pair (a, b).

(2). The inequality in equation (8) may be written as

$$\frac{1}{2\pi \sin A \sin B \sin C}\left|\frac{1}{\sqrt{1-\left(\frac{2}{k_i p_i}\right)^2}}\right|$$

$$\geq \frac{1}{2\pi \sin A \sin B \sin C}\left|\frac{1}{\sqrt{1-\left(\frac{2}{k_j p_j}\right)^2}}-1\right|$$

which reduces to the following inequality

$$k_i p_i \leq k_j p_j$$

The kp term using the l and k parameters (in section 2) is

$$kp = k\left(\frac{1+l^2+k^2}{k^2}\right) = k + \frac{1}{k} + \frac{l^2}{k}$$

Now the minimum value of this expression (since k is always >0) is when $k=1$ and $l=0$, which occurs when $A=$a and $B=$b i.e. when the image angles equal the model angles. The decision rule is therefore

$$Select\ \bar{T} = T_i\ if\ k_i + \frac{1}{k_i} + \frac{l_i^2}{k_i} \leq k_j + \frac{1}{k_j} + \frac{l_j^2}{k_j} \quad (9)$$

4.2 No Correspondence

When there is no knowledge about which vertex in an image corresponds to which vertex in a model triangle, then the decision rule in equation (9) changes to:

$$\bar{T} = T_i\ if\ f(A, B|T_i) \geq f(A, B|T_j)\ \forall j \neq i \quad (10)$$

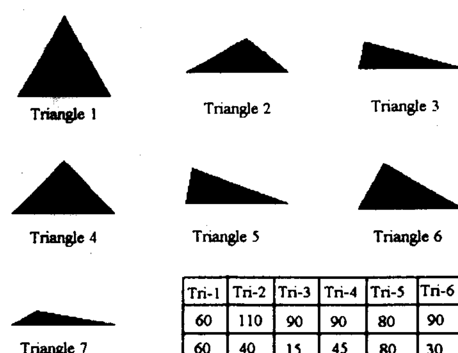
where the density is the triangle angle-pair density given in equation (2).

5. Experiments

The usefulness of the derived probability density function was tested in a series of experiments and reported in the previous paper. Various samples as shown in (Fig. 7) were selected and imaged on the target structure sketched in (Fig. 8). The samples used in the experiments are labeled and described in ⟨Table 2⟩.

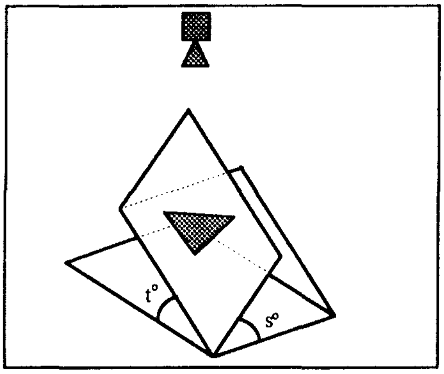
The vertices in the images in (Fig. 9) were extracted and the angles computed (see ⟨Table 3⟩). For all model triangles we assume there is no angle correspondence, thus equation (10) is used for the decision rule. A tabulation of all these values is shown in ⟨Table 4⟩.

Consider the Sample-a values shown in the second colum of the table. The likelihood function values

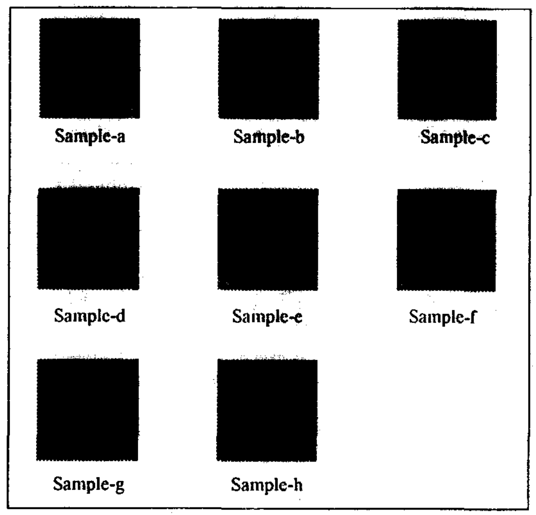


Tri-1	Tri-2	Tri-3	Tri-4	Tri-5	Tri-6	Tri-7
60	110	90	90	80	90	140
60	40	15	45	80	30	10
60	30	75	45	20	60	30

(Fig. 7) Model triangles and their angles



(Fig. 8) A sketch of the target area showing the target *orientation*



(Fig. 9) Images of sample triangles after certain orientations

〈Table 2〉 Sample triangles and their approximate orientations.

Sample	Triangle	t°	s°
a	2	16	40
b	1	38	44
c	3	35	40
d	4	38	45
e	7	35	43
f	1	40	55
g	5	21	32
h	1	21	40

〈Table 3〉 Measured angles of sample triangles after certain *orientations*.

Sample	Measured		
a	31.6770	101.8908	46.4322
b	73.0989	55.3940	51.5071
c	63.0794	99.9854	16.9352
d	45.6658	49.0547	85.2795
e	24.1325	145.3833	10.4842
f	54.0058	82.2607	43.7335
g	73.0507	84.2267	22.7225
h	69.1531	55.2777	55.5693

〈Table 4〉 Likelihood values for different samples.

Triangle	Samp-a	Samp-b	Samp-c	Samp-d	Samp-e	Samp-f	Samp-g	Samp-h
1	0.2250	**1.4287**	0.0665	0.6114	0.0199	0.6623	0.1311	**2.2452**
2	**0.8447**	0.1874	0.0479	0.2237	0.0196	0.1895	0.0878	0.1769
3	0.0513	0.0532	**0.4584**	0.0531	0.0264	0.0564	0.2457	0.0528
4	0.3391	0.5833	0.0664	**1.5845**	0.0235	**0.7597**	0.1236	0.5215
5	0.0990	0.1003	0.3132	0.0977	0.0247	0.1043	**0.8004**	0.0994
6	0.1467	0.2479	0.1511	0.2737	0.0356	0.3291	0.3135	0.2497
7	0.0172	0.0201	0.0747	0.0222	**0.2288**	0.0233	0.0404	0.0200
Decision	**2**	**1**	**3**	**4**	**7**	**4**	**5**	**1**

evaluated using the seven different triangles are shown. The largest value of the likelihood function in this case is 0.8447 and corresponds to triangle 2. The next most likely in this case is triangle 4 with a likelihood value of 0.3391. The decision, of course, is triangle 2. We can see from table 2 that this is indeed the correct decision. In fact the scheme gave incorrect results only in the case of Sample-f. It should be remembered that the method is statistical and will not always give the correct result. The method becomes less reliable, as is to be expected the sharper the target orientation angles (t and s).

6. Conclusions

The joint density of the two measured angles in an orthographic projection has been derived and extended to the joint density of the angles in any polygon. Using this density function optimum decision rule for recognizing geometric shapes in images has been proposed and demonstrated its use to distinguish amongst the triangles in a group. From the experiment conducted, the joint density function turns out to be a quite reliable method. The method is not limited to two-dimensional objects because a triangle captures the spatial arrangement of identifiable landmarks in any three dimensional scene. The importance of being able to distinguish similar shapes arises from the fact that surface segments (such as planar faces) of distinct objects are often similar.

This method, however, needs to be extended to model the errors that arise from incorrect positioning of landmarks in an image.

References

[1] P.J. Besl and R. C. Jain, "Three-dimensional object recognition," ACM comput. Surveys, vol. 17, no. 1, pp. 75-145, 1985.

[2] R. T. Chin and C. R. Dyer, "Model-based recognition in robot vision," ACM Comput. Surveys, vol. 18, no. 1, pp. 67-108, 1986.

[3] P. Suetens, P. Fua, A.J. Hanson, "Computational Strategies for Object Recognition," ACM Computing Surveys, Vol. 24, No. 1, pp. 5-61, 1992.

[4] J.B. Burns, R.S. Weiss, E.M. Riseman, "View Variation of Point-Set and Line-Segment Features," IEEE Trans. Pattern analysis and Machine Intelligence, Vol. 15, No. 1, pp. 51-68, 1993.

[5] J. Ben-Arie, "The Probabilistic Peaking Effect of Viewed Angles and Distances with Application to 3-D Object Recognition," IEEE Tran. on Pattern Analysis and Machine Intelligence, Vol. 12, No. 8, pp. 760-774, 1990.

[6] T.O. Binford, T.S. Levitt, and W.B. Mann, "Bayesian Inference in Model Based Machine Vision," L. Kanal, T. Levitt, and J. Lemmer, eds., Uncertainty in Artificial Intelligence 3. New York : Elsevier Science Publishers B.V. (North Holland), 1989.

[7] R. Malik, T. Whangbo, "Junction View Densities in Images," Proc. of IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 142-147, 1994.

[8] R. Malik, T. Whangbo, "View Variations in Angles," Proc. of Third GWIC on Intelligence Systems, June 1994.

[9] R. Malik, T. Whangbo, "Using Angle Densities for 3-D Recognition," Proc. of SPIE Internal Symposium on Intelligent Robots and Computer Vision XIII, November 1994.

[10] R. Malik, T. Whangbo, "Angle Densities and Recognition of 3D Objects," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 1, pp. 52-57, 1997.

[11] T. Whangbo, "3-D Object Recognition Using Angle View Densities," PhD dissertation, Dept. of EECS, Stevens Inst. Of Tech., May 1995.

[12] A. Papoulis, "Probability, Random Variables, and stochastic Processes," McGraw-Hill Inc., 1965.

황 보 택 근

1983년 고려대학교 금속공학과 졸업(학사)
1987년 뉴욕시립대학교 전자계산학과(공학석사)
1995년 Stevens Inst. Of Tech. 전자계산학과(공학박사)
1988년~1992년 Q-Systems Technical Manager
1995년~1997년 삼성종합기술원 선임연구원
1997년~현재 경원대학교 전자계산학과 전임강사
관심분야: Computer Vision, Computer Graphics