A MODEL BASED VISION SYSTEM FOR RECOGNITION OF MACHINE PARTS

Katsushi Ikeuchi
Yoshiaki Shirai

Computer Vision Section
Electrotechnical Laboratory
Umezono-1-1-4, Sakura-mura
Niihari-gun, Ibaraki, 305
Japan

ABSTRACT

This paper describes a vision system based on the photometric stereo system and a model generator called GEOMAP. The photometric stereo system obtains a needle map from shading information of an observed image. An extended Gaussian image of the needle map reduces possible attitudes of an object relative to the viewer. A model-generator called GEOMAP generates a needle map which would be observed from the viewer direction determined from EGI. Comparing the needle map by the GEOMAP with the needle map by the photometric stereo system makes final recognition of the object.

I INTRODUCTION

Machine vision at the low level often provides a collection of local surface normals [1,2,3]. This collection of local surface normals is refered to as a needle map [1]. In particular, the photometric stereo system [4,5,6,7] can provide a needle map from shading information.

A global recognition of an object requires interpreting these local representations. Each local representation is obtained at a particular point in the viewer-centered coordinate system. On the other hand, an object model is usually expressed in a coordinate system based on an object center and natural axis of the object [8,9]. These two coordinate systems are independent of each other. Thus, we have to recover the object-centered coordinate system from local representations on the viewer-centered coordinate system.

We propose to use the extended Gaussian image (EGI) [1,10,11,12,13] as a constraint to reduce possible attitudes of an object relative to the viewer. Once an attitude is determined, we can easily compare a representation on the viewer-centered coordinate system with a representation on the object-centered coordinate system.

GEOMAP [14] can generate a needle map which would be observed from the viewer direction determined from EGI. Fig. 1 denotes a total schema of our vision system. Photometric stereo system provides a needle map from three brightness arrays. An extended Gaussian image from the needle map determines an object's attitude. GEOMAP generates a needle map based on the obtained attitude. Comparison between the observed needle map and the generated needle map gives a final decision on what kind of an object is observed.

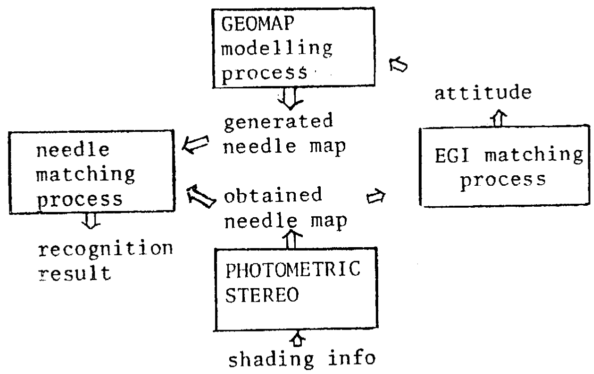


Fig. 1 Total schema

II PHOTOMETRIC STEREO SYSTEM

Photometric stereo system obtains needle maps from shading information [4,5,6,7]. The orientation of patches on the surface of an object can be determined from multiple images taken with different illumination, but from the same viewing direction. The method is refered to as photometric stereo method [4]. The photometric stereo can be implemented using a lookup table based on numerical inversion of reflectance map[1,6]. A needle map is obtained from three brightness images using the lookup table.

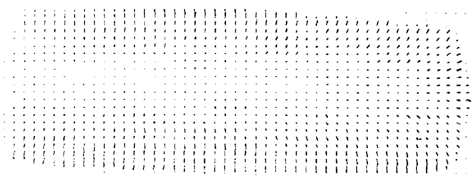
The brightness distribution of a sphere under the lighting system provides a reflectance map [5,6]. Since we know surface orientation of a sphere at each image point, we can obtain a reflectance map which denotes the brightness value at each surface orientation. We use a Lambertian sphere to obtain a reflectance map for diffuselly reflecting surfaces under the point source illumination.

Each image point provides three brightness values ($I1,I2,I3$). Then they are normalized as

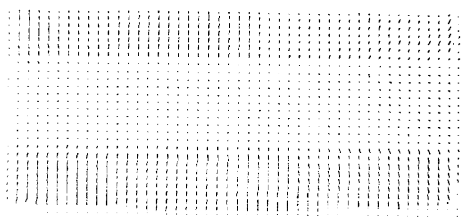
$$E_i=I_i/(I1+I2+I3),\ i=1,2,3$$

to cancel the effect of albedo and shading effect of a lens. This operation provides three reflectance maps. Three reflectance maps are numerically inverted into a lookup table which gives surface orientation from three brightness values.

Fig 2 shows an example of needle maps obtained from real objects (a cylinder, an octagonal prism, and a square prism) using the system. A TV camera looks down the objects from the ceiling. The objects lie having their long axes along the x coordinate of the image plane. This looking-down-tv method gives a good constraint on the possible viewer directions as discussed in the next section. Since these three objects have rectangular silhouette in this attitude, three objects cannot be discriminated using the so-called silhouette matching.



(a) Cylinder



(b) Octagonal prism

(c) Square prism

Fig. 2 Obtained needle maps from the photometric stereo

III DETERMINING ATTITUDE OF AN OBJECT

There are three degrees of freedom in an object's attitude relative to the viewer. Although a brute force technique, such as search through the space of possible attitudes, can be applicable to matching of EGIs directly for determining attitude of an object, we will reduce this search space using constraints.

A. Constraints on attitudes of an object

We propose three kinds of constraints to reduce possible attitudes of an object. (From now on, we assume that image plane is perpendicular to the line of sight. In other words, the image formation system is modelled using the orthographic projection.)

Knowing the gravity direction reduces possible attitudes. Two conditions must be satisfied to apply this constraint to the observed data:

1) Stable attitudes of models must be known.
   If we have a solid model, we can determine the position of gravity center. In particular, GEOMAP has a solid model, and can determine the position of gravity center. Then, we can obtain stable attitudes of the model from the position of gravity center. We denote the opposite direction to the gravity at the stable attitude as <u>the stability direction of the model</u>.
2) The angle between the gravity direction and the line of sight must be detected.
   It is interesting that a human being also has semicircular canals to detect the direction.

Let us assume that we can detect the angle between the gravity direction and the line of sight from a level on a TV camera as ω . The gravity direction is parallel to the stability direction from the definition. Also the line of sight has always the same direction from the assumption of orthographic projection. Thus the angle between the stability direction and the line of sight has the same angle ω . In other words, the possible line of sight direction locates on a circle on the Gaussian sphere whose center corresponds to the stability direction and whose radius corresponds to the angle between the gravity direction and the line of sight, ω . In particular, when TV camera observes the object in the gravity direction as in our case, the circle reduces to the center point. For example, when a square prism is looked down from the ceiling of the observer room, only six points corresponding to prism's faces on the Gaussian sphere are necessary to be considered as candidate directions of line of sight. This constraint also works under the partial observations.

The ratio of area projected onto the image plane against the original surface area constrains possible line of sight direction [11]. For example, observing an ellipsoid from the long-axis direction gives a smaller projected area than looking at the same ellipsoid from the direction perpendicular to the axis. Yet the surface area is the same.

We also must find the rotation of an object around the line of sight. We use the principal EGI inertia direction from the line of sight direction [11]. We will use the 2-D axis for simplicity of calculation, where the calculations are also done only if surface patch is visible from the direction.

The constraints mentioned above can be extended to treat partial observations. Partial observation occurs due to characteristics of photometric stereo system. The photometric stereo system can determine surface orientation at the area where all three light sources project their light directly. Our photometric stereo system can determine surface orientation within the area where the zenith angle of the surface normal is less than 60 degrees. The ratio and the axis direction are calculated at this area.

B. Determining attitude using EGI

Comparing EGI from observed data with EGI from prototypes determines the attitude of an object. An EGI of an object denotes surface-normal distribution of the object [1,10,11,12,13]. We implement EGI on the semi-regular geodesic dome from a two frequency dodecahedron [11]. EGI at each attitude is calculated using GEOMAP. EGI is stored as a two dimensional lookup table. Each column corresponds to one attitude, while each raw contains two kinds of information.

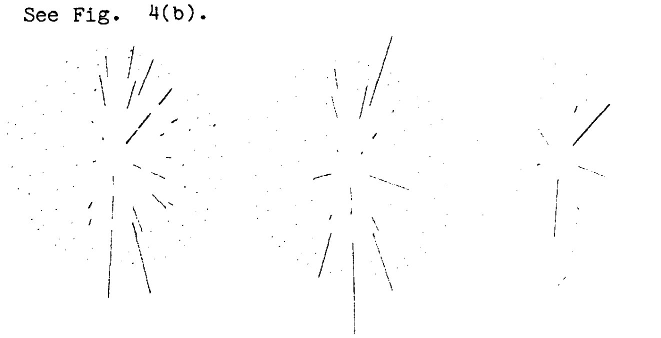
(1) projected area vs. surface area.
Comparing this registered ratio with an observed ratio determines whether the EGI distribution is necessary to be check at this attitude or not.

(2) local EGI distribution
An EGI observed at this attitude is stored in an one-dimensional table based on the geodesic-dome-tessellation. The EGI in the table is aligned so that the least inertia axis coincides with the x-coordinate on the image plane. Thus, the observed EGI should be rotated to have the least inertia axis as the direction before comparison.

Likelihood is calculated by determining whether every observed EGI cell contains similar amount of mass at the corresponding neighboring cells in the table [11]. The program determines the attitude obtaining the highest likelihood as the observed attitude.

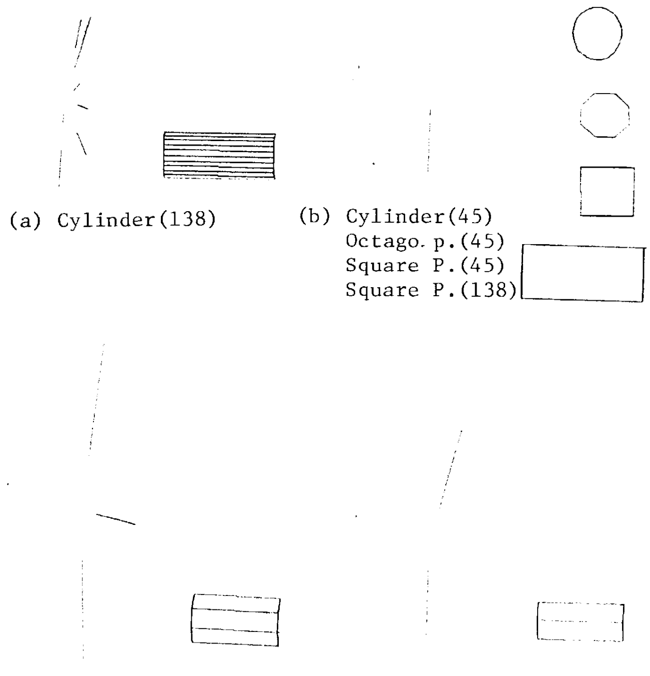
Table 1 shows a EGI comparison results. A needle map of a cylinder (Fig.2(a)) gives an EGI as shown in Fig.3(a) and 0.8711 as the ratio. If the needle map is assumed to come from a cylinder, the ratio constraint reduces 240 possible attitudes into 2 possible attitudes: direction 22 and 138. See Fig. 4. At 22 direction, EGI comparison gives 0.3437 as likelihood value. If the same needle map is assumed to come from a octagonal prisms, 22, 136 and 138 are possible attitudes. Under this assumption, direction 138 gives the highest likelihood among the possible attitudes. At the EGI comparison the observed EGI is too different to be assumed as an EGI from the square prism.

The needle map in Fig. 3(c) comes from the planar surface of the prism. All objects have planar surfaces. The direction 45 corresponds to the direction observing the base of each prism. See Fig. 4(b).



(a) Cylinder (b) Octagonal prism (c) Square prism

Fig.3 Obtained EGIs from the needle maps



(a) Cylinder(138) (b) Cylinder(45) Octago. p.(45) Square P.(45) Square P.(138)

(c) Octagonal P. (138) (d) Square P.(22)

Fig. 4 Prototypes and their EGIs

Table 1 EGI comparison results

Candidate / Object exposed	CYLINDER	OCTAGONAL PRISM	SQUARE PRISM
CYLINDER (0.8711)	22: 0.3305 138: 0.3437	22: 0.0 136: 0.0 138: 0.0998	22: 0.0 136: 0.0
OCTAGONAL PRISM (0.8766)	22: 0.3137 138: 0.3494	22: 0.0 136: 0.0 138: 0.1250	22: 0.0 136: 0.0
SQUARE PRISM (0.9803)	45: 0.2774	22: 0.2483 45: 0.2774 136: 0.2531 159: 0.0840	22: 0.1338 45: 0.2774 136: 0.0 138: 0.2774 159: 0.0840

IV GEOMAP

GEOMAP, developed by Kimura and Hosaka at ETL [14], is a program package to synthesize 3D object. GEOMAP maintains 3D information of an object in the program. GEOMAP has the following characteristics;

(0) GEOMAP has primitive objects.
(1) GEOMAP can move and rotate an object.
(2) GEOMAP can unify an object with another object into a new object.
(3) GEOMAP can make constraint endowment.
(4) GEOMAP can make a 2D projection of an object in any direction.
(5) GEOMAP can store/retrieve data of each object to/from files.
(6) Maclisp at ETL can call GEOMAP directly with commands to GEOMAP.

Our vision system is assumed to treat mechanical parts. We have precise information at each candidate object. Thus, we can make 3D information of each candidate in GEOMAP using movement operations, rotation operations, and unification operations from primitive objects.

Once the EGI system determines the attitude of each candidate, GEOMAP can make a needle map of each candidate in the direction. Needle map comparison has two relative merits to a depth map, which would be obtained by integrating the needle map, comparison.

(1) The photometric stereo system does not assume that an object surface is continuous. Thus, we cannot integrate a needle map to obtain a depth map. Even if we assume that an observed surface is continuous, we cannot obtain a depth map in case that needles are obtained at multiple regions.

(2) Each needle may contain a measurement error. Thus, errors accumulate during integration. Reliability of a depth map degenerates towards the end of integration.

GEOMAP generates a needle map so as to have the same gravity position, to have the least inertia direction as the x-coordinate axis, and to have the same window size as the real image. Final matching is made as

$$E = \frac{\Sigma_i \left(1 - \frac{\sqrt{(f_i^{proto}-f_i^{observe})^2+(g_i^{proto}-g_i^{observe})^2}}{2} \right)}{\text{window-area}}$$

After calculating this E for all candidates, the candidate obtaining the highest value will be determined as the object observed at that time as shown in Table 2.

Table 2 Needle-comparison results

Object exposed \ Candidate	CYLINDER	OCTAGONAL PRISM	SQUARE PRISM
CYLINDER	183: 0.7834	183: 0.5859	NIL
OCTAGONAL PRISM	183: 0.7365	138: 0.7506	NIL
SQUARE PRISM	45: 0.2649	45: 0.2711	45: 0.3212 138: 0.8086

V CONCLUSION

We propose a vision system based on the photometric stereo system and GEOMAP. The photometric stereo system provides a needle map from shading information. After constraints from EGI reduces possible attitudes of an object in space, GEOMAP generates a needle map of an object at the attitude. Comparing the observed needle map with the generated needle map identifies the observed object.

The experiment shows that extending the detectable area of a photometric stereo system not only increases accuracy in final matching process but also constrains possible attitudes more precisely.

ACKNOWLEDGMENTS

The authors would like to extend their sincere appreciation to Prof. B.K.P. Horn of MIT for his helpful discussions. Thanks goes to Prof. F. Kimura of Univ.of Tokyo for permission of using GEOMAP. GEOMAP is converted from VAX-FORTRAN to TOPS20-FORTRAN. The experimental part of this paper was done using Mac-Lisp installed on Dec 2020-2060 at ETL.

REFERENCES

[1] Horn, B.K.P., "SEQUINS and QUILLS -- Representation for Surface Topography." AI Memo No. 536, AI Lab., MIT, 1979.

[2] Barrow, H.G. and Tenenbaum, J.M., "Recovering Intrinsic Scene Characteristics from Images." Hanson, A.R. & Riseman, E.M., (Ed.) Computer Vision Systems, Academic Press, New York, 1978.

[3] Kender, J.R., "Shape from Texture," Ph.D. Thesis, Computer Science Dept., CMU, 1980.

[4] Woodham, R.J., "Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from Image Intensity," in Proc. SPIE, Vol.155, 1978.

[5] Ikeuchi, K., "Determining Surface Orientations of Specular Surfaces by Using the Photometric Stereo Method," IEEE Trans. on PAMI, Vol.PAMI-2, No.6, Nov. 1981, 661-669.

[6] Silver, W.M., "Determining Shape and Reflectance Using Multiple Images," MS Thesis, EECS, MIT, June 1980.

[7] Coleman, E.N., and Jain, R.,"Shape form Shading for Surfaces with Texture and Specularity," in Proc. IJCAI-81, Vancouver, Aug. 1981, 652-657.

[8] Binford, T.O.,"Visual Perception by Computer." IEEE Conf. on Systems and Control, Miami, Dec. 1971.

[9] Marr, D. and Nishihara, K.H., "Representation and Recognition of the Spatial Organization of Three-dimensional Shape." Proc. R. Soc. Lond., B:200, (1978).

[10] Smith, D.A.,"Using Enhanced Spherical Images for Object Representation," AI Memo No. 530, AI Lab., MIT, 1979.

[11] Ikeuchi,K.,"Recognition of Objects Using Extended Gaussian Image'" in Proc. IJCAI-81, Vancouver, Aug. 1981, 595-600.

[12] Ballard, D.H., and Sabbah, D.,"On Shape," in Proc. IJCAI-81, Vancouver, Aug. 1981, 607-612.

[13] Bajcsy, R.,"Three-dimensional Scene Analysis," in Proc. ICPR-80, Miani Beach, Dec. 1980, 1064-1074.

[14] Hosaka, M, Kimura, F. and Kakishita, N., "A Unified Method for Processing Polyhedra," in Proc. IFIP, 1974, 768-772.