'91 KACC 1991. 10. 22~24

Recognizing a Polyhedron by Network Constraint Analysis

Seiji Ishikawa, Mayumi Kubota, Hiroshi Nishimura, Kiyoshi Kato

Department of Electrical, Electronic and Computer Engineering
Kyushu Institute of Technology
Tobata, Kitakyushu 804 JAPAN

Abstract The present paper describes a method of recognizing a polyhedron employing the notion of network constraint analysis. Typical difficulties in three-dimensional object recognition, other than shading, reflection, and hidden line problems, include the case where appearances of an object vary according to observation points and the case where an object to be recognized is occluded by other objects placed in its front, resulting in incomplete information on the object shape. These difficulties can, however, be solved to a large extent, by taking account of certain local constraints defined on a polyhedral shape. The present paper assumes a model-based vision employing an appearance-oriented model of a polyhedron which is provided by placing it at the origin of a large sphere and observing it from various positions on the surface of the sphere. The model is actually represented by the sets of adjacent faces pairs of the polyhedron observed from those positions. Since the shape of a projected face gives constraint to that of its adjacent face, this results in a local constraint relation between these faces. Each projected face of an unknown polyhedron on an acquired image is examined its match with those faces in the model, producing network constraint relations between faces in the image and faces in the model. Taking adjacency of faces into consideration, these network constraint relations are analyzed. And if the analysis finally provides a solution telling existence of one to one match of the faces between the unknown polyhedron and the model, the unknown polyhedron is understood to be one of those memorized models placed in a certain posture.

In the performed experiment, a polyhedron was observed from 320 regularly arranged points on a sphere to provide its appearance model and a polyhedron with arbitrarily postured, occluded, or imposed another difficulty was successfully recognized.

1. INTRODUCTION

Understanding scenes with various objects has been vigorously investigated in the field of robot vision and many techniques have been proposed including description, reconstruction and recognition of objects([1], [2]). Yet the difficulties peculiar to three-dimensional scenes have not been solved in a satisfactory manner. Typical difficulties in three-dimensional object recognition, other than shading, reflection, and hidden line problems, include the case where appearances of an object vary according to observation points and the case where an object to be recognized is occluded by other objects placed in its front, resulting in incomplete information on its shape.

What should be considered first to overcome these difficulties is how to describe the objects concerned. Minsky[3] describes various appearances of an object employing a network whose end nodes represent the features of projected shapes of the object, while Koenderik and Van Doorn[4] propose an aspect graph which contains all the topologically different shapes of an object obtained as the result of an observer's migration around it. These models are often referred to as appearance oriented models and the model also includes Fekete and Davis's study[5] where they propose a property sphere. These descriptions are, however, unlikely to be directly applicable to occluded cases.

In the second place, a recognition technique should be chosen carefully to solve these difficulties. Since any two-dimensional image of an object will never give information on its entire shape, it is reasonable to focus our attention on the local geometrical constraints of an object and to employ the technique called a network constraint analysis. One of the earliest researches that employ the notion of local constraints is done by Cherry and Vaswani[6] in which a fast technique is shown of assigning binary values to logical variables in a given set of logical formulae. Waltz[7] employs this concept in scene labeling to realize fast correspondence between geometrical labels and the vertices on a polyhedral line drawing. His technique has been extended to a relaxation operation by Rosenfeld *et al.*[8], which is an established technique in scene analysis. These constraint analysis problems have been generalized and formalized as a consistent labeling problem by Haralick and Shapiro[9].

The technique proposed in the present paper for solving the above mentioned difficulties in three-dimensional object recognition makes use of the geometrical constraints given by the appearances of adjacent faces on an object concerned. All the possible appearances of each adjacent faces pair of an object are stored into memory and used to find corresponding faces on an acquired image. The entire matching process can be described as a network constraint analysis. This technique is elaborated in the following along with some experimental results.

2. DESCRIBING POLYHEDRAL SHAPE

Here the idea is viewed of the constraints among adjacent faces on an object. Let us take a cube as an example. Each face on a cube has the shape of a square and any adjacent faces pair meet perpendicularly one another. Now if a cube is projected on a plane and a projected face, say f_1, has the shape shown in **Fig.1**a, two of its adjacent faces, say f_2 and f_3, should have respective shapes such as shown in Fig.1a. In the same way, if face f_1 has the shape given in Fig.1b, faces f_2 and f_3 must be those given in Fig.1b. These three faces therefore constrain each other geometrically when projected on a plane. The matching technique proposed in the present paper employs this kind of constraints and is applied to polyhedral recognition.

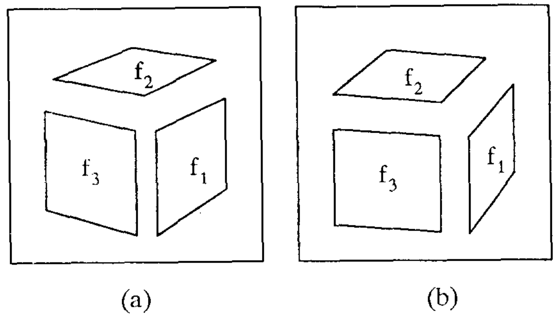


Fig.1 Projection of the faces on a cube onto a plane.

Since two-dimensional shapes are employed for the matching, there is no need of three-dimensional reconstruction of the object concerned for recognition. Hence we assume a single view. Obviously a large amount of reference data (actually every conceivable projection) need to be prepared for finding match. For convenience, we restrict the objects in a scene to polyhedra in this particular paper. Application of the present technique to curved objects is, however, within the range of this research. The projection is assumed to be orthographic. Perspective projection simply enlarges the amount of the reference data.

Let us denote a polyhedron by P, the faces on P by f_i (i=1,2,..., n_P), and the edges and the vertices on face f_i by e_{ij} and v_{ij} (j=1,2,..., n_i), respectively. Let us also denote the length of edge e_{ij} on a plane projected from the kth direction by $d_{ij}(k)$, and an inner angle at vertex v_{ij} on the projected plane by $\theta_{ij}(k)$. Then a face f_i projected from the kth direction (denoted by $p_k(f_i)$) is expressed by a set of the values in the following way:

$$p_k(f_i) \equiv (d_{i1}(k),\ \theta_{i1}(k),\ d_{i2}(k),\ \theta_{i2}(k),\ \ldots\ldots,\ d_{in_i}(k),\ \theta_{in_i}(k)) \qquad (1)$$

These components are chosen clockwise on a projected face. Practically a normalized value $s_{ij}(k)=d_{ij}(k)/\max\{d_{ij}(k), j=1,2,\ldots,n_i\}$ is employed for $d_{ij}(k)$.

Since the proposed matching technique employs appearances of each face on a polyhedron, the reference data for the matching are to include in principle the shapes of each face projected from every conceivable direction. This is practically provided by placing a polyhedron at the origin of a large sphere and observing the shape of each face from various positions on the surface of the sphere. The sphere comes from what was formerly proposed by Fekete and Davis[5] as a property sphere, where some geometrical features were observed and registered at every arranged point on the surface of a sphere. Observation points on a sphere are provided, in the simplest case, as the centroid of each triangular face of a regular icosahedron, *i.e.*, 20 observation points. As each triangle can be separated into four smaller triangles by connecting the middle points of its sides, observation points can be increased by the factor of four times. **Figure 2** illustrates some appearances of a polyhedron at observation points on a sphere. Using this, all conceivable appearances are stored into memory with respect to each pair of adjacent faces on a polyhedron. If, when face f_i has the appearance(projected shape) $p_k(f_i)$, adjacent face f_j has the appearance $p_k(f_j)$, then 4-tuple $(f_i, p_k(f_i), f_j, p_k(f_j))$ is stored into memory. Let us

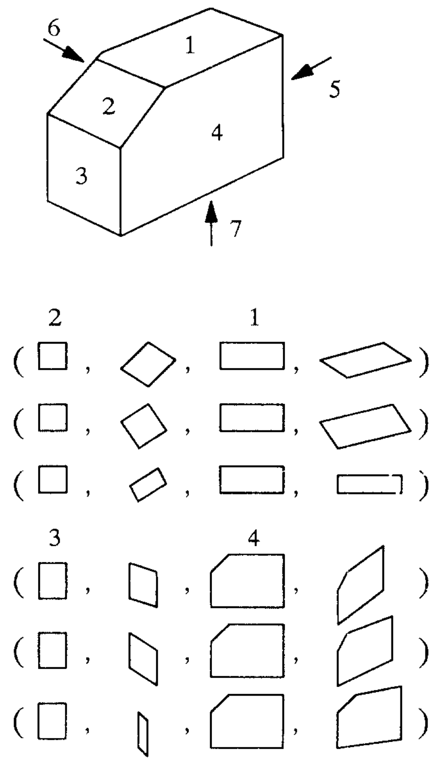


Fig.3 The appearance-oriented knowledge with respect to a polyhedron.

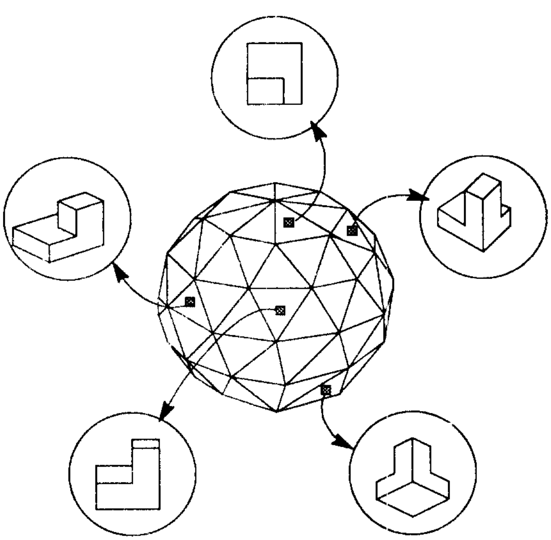


Fig.2 Obtaining appearances of a polyhedron from the observation points on a sphere.

denote the set of all adjacent faces pairs on a polyhedron P by T and the set of the above 4-tuples by R. Then

$$T=\{(f_i, f_j) \mid adj(f_i, f_j)=1\},$$
$$R=\{(f_i, p_k(f_i), f_j, p_k(f_j)) \mid (f_i, f_j)\in T,\ v_k\in V\}.$$

Here adj(f, g)=1, if faces f and g are mutually adjacent, and 0, otherwise. The set V includes all the visual points v_k (k=1,2,...,n_v) on the sphere. The set R is the knowledge on the model polyhedra the present recognition technique holds. Some elements of R is shown in **Fig.3**.

3. PROCEDURE FOR RECOGNITION

The procedure for recognizing a polyhedron is first to find correspondences between faces on an acquired polyhedral image and those in the reference data(R), and then to eliminate incorrect correspondences employing the adjacency relation among the faces. This process can be referred to as a network constraint analysis problem.

Given an observation point v_k, an image of an unknown polyhedron P' is obtained. Assume that a visible face f_u' on P' corresponds to an area a_u (u=1,2,...,$n_{P'}$) on the image. Note that $a_u \equiv p_k(f_u')$. The correspondence between these areas and those memorized faces in R is checked and the following renewal of the elements of R is performed by **CORRESPOND** shown in **Fig.4** to memorize the correspondence. For $\mathbf{r}=(f_i, p_k(f_i), f_j, p_k(f_j)) \in R$;

1. $\exists u, \exists v, p_k(f_i)=a_u, p_k(f_j)=a_v$
   $\mathbf{r} \rightarrow (u, p_k(f_i), v, p_k(f_j))$
2. $\exists u, p_k(f_i)=a_u$
   $\mathbf{r} \rightarrow (u, p_k(f_i), nil, p_k(f_j))$
3. $\exists v, p_k(f_j)=a_v$
   $\mathbf{r} \rightarrow (nil, p_k(f_i), v, p_k(f_j))$

```
CORRESPOND
begin
  copy R to R*;
  for all r =(fi, pk(fi), fj, pk(fj)) ∈ R* do
    begin
      flag0:=0;
      flag1:=0;
      for all u do
        begin
          if pk(fi)=au
            then begin
                    replace r1 by u;
                    flag0:=1
                 end;
          if pk(fj)=au
            then begin
                    replace r3 by u;
                    flag1:=1
                 end
        end;
      if flag0=0
        then if flag1=0
               then R*:=R*-{ r }
               else replace fi by nil
        else if flag1=0
               then replace fj by nil
    end
end.
```

Fig.4 Procedure **CORRESPOND** which finds possible correspondences between the projected faces on a polyhedral image and those in R. Note that $(f_i, p_k(f_i), f_j, p_k(f_j)) \equiv (r_1, r_2, r_3, r_4)$.

```
FILTER
begin
  for all the possible assignments (u, pk(fi))
                         with which R* provides do
    if for all t'=(pk(fi), pk(fj))∈T'(pk(fi))
          there exists r ∈R* such that
          r=(u, pk(fi), v, pk(fj)) and adj(u, v)=1
      then retain the assignment pk(fi) to u
      else replace u by nil;
  for all r=(u, pk(fi), v, pk(fj)) ∈ R* do
    if u=nil or v=nil
      then discard r from R*
end.
```

Fig.5 Algorithm **FILTER** which examines true match.

Here u,v =1,2,...,$n_{P'}$. These renewed 4-tuples are stored into the set R^*.

Elimination of incorrect correspondences with respect to the face $p_k(f_i)$ in R^* is done employing the set $T'(p_k(f_i))$ defined by

$$T'(p_k(f_i))=\{(p_k(f_i), p_k(f_j)) \mid adj(p_k(f_i), p_k(f_j))=1\}.$$

This is performed by **FILTER** shown in **Fig.5**, through which the elements of the set R^* receive discarding, if any, to give a final result of the recognition. One may understand that a polyhedron on an image has been recognized as one of the model polyhedra, if R^* reports one to one match between those faces on the image and the model.

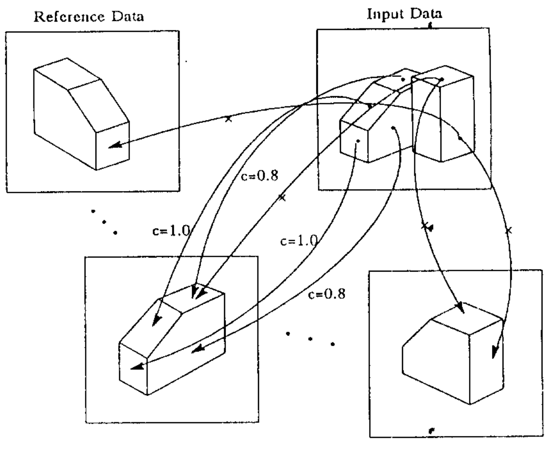
Since the present research also deals with occluded cases, a certainty measure c is introduced with respect to the degree of match between faces in the following way;

1) If two faces agree entirely with their representations(1), then c=1.
2) If two faces agree with three successive angles and with the lengths of the edges between them, then c=0.75.
3) If two faces agree with two successive angles, then c=0.5.

The certainty measure of match between faces f_i and f_j is denoted by c_{ij}. The entire certainty measure C is defined by the following formula;

$$C = \sqrt[p]{\prod_{i=1}^{p} c_{ij}} \quad ,$$

where p is the number of the faces on the acquired polyhedral image. **Figure 6** illustrates an overall idea of the recognition process.



C=0.89

Fig. 6. Illustration of correspondences between the areas on an acquired image and the faces in *R*, and of retaining/discarding the correspondence according to the geometrical constraints provided by adjacent faces. Also shown is a certainty measure. Note that, although 4-tuples in *R* should be illustrated in the above figure as reference data, appearances of whole polyhedral shape are shown for convenience.

4. EXPERIMENTAL RESULTS

An experiment is performed employing two white polyhedra one of which is shown in **Fig.7** with face labels. The sphere employed to prepare for the reference data *R* has 320 observation points. An unknown polyhedron is taken an image by a CCD camera and it is transformed into a 256×256 pixels digital image with 64 gray levels to process by a 16-bit personal computer. Images of a polyhedron in several situations are taken and, after preprocessing, each area corresponding to a face of unknown polyhedron is temporally given labels h_j. The proposed recognition technique is applied to these labeled areas and the correspondence between those areas(h_j) on the image and faces(f_i) in R^* is examined.

Figure 8 shows some results: Fig.8a–c are the cases where the polyhedron in Fig.7 has various postures; and Fig.8d realizes an occluded case. With all cases, the polyhedron was successfully recognized which one can see in the respective matrix in Fig.8 indicating the correspondence between faces h_j and faces f_i.

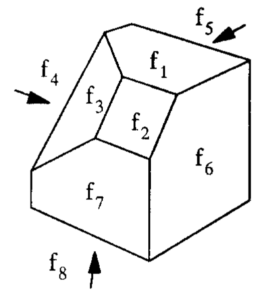
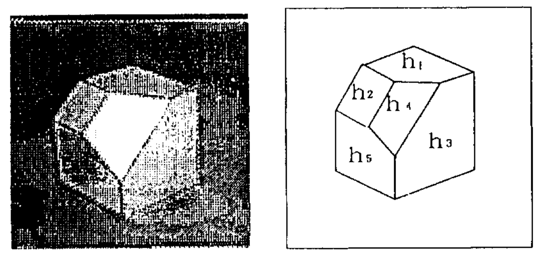


Fig.7 One of the polyhedra employed in the experiment.



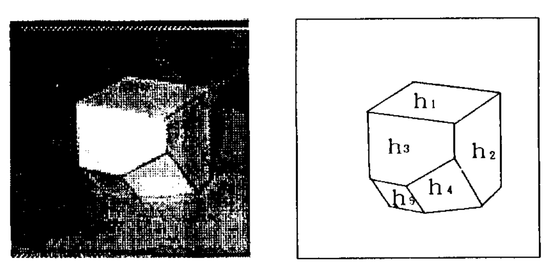
an input image visible faces

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
h_1	0	0	0	0	9	9	1	9
h_2	0	1	0	0	9	9	0	9
h_3	0	0	0	1	9	9	0	9
h_4	0	0	1	0	9	9	0	9
h_5	1	0	0	0	9	9	0	9

C=1.000

result of the recognition

(a)

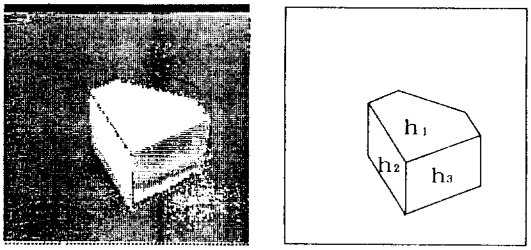


	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
h_1	9	0	0	0	9	9	0	1
h_2	9	0	0	1	9	9	0	0
h_3	9	0	0	0	9	9	1	0
h_4	9	0	1	0	9	9	0	0
h_5	9	1	0	0	9	9	0	0

C=1.000

(b)

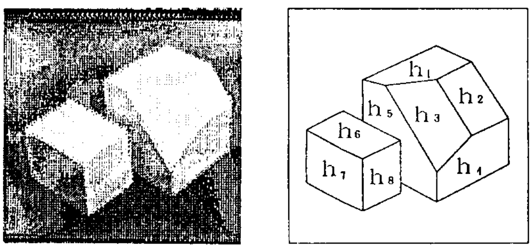
Fig.8 Experimental results.



	f1	f2	f3	f4	f5	f6	f7	f8
h1	9	9	9	1	0	9	9	0
h2	9	9	9	0	0	9	9	1
h3	9	9	9	0	1	9	9	0

C=1.000

(c)



	f1	f2	f3	f4	f5	f6	f7	f8
h1	1	0	0	0	9	9	0	9
h2	0	1	0	0	9	9	0	9
h3	0	0	1	0	9	9	0	9
h4	0	0	0	0	9	9	1	9
h5	0	0	0	2	9	9	0	9
h6	0	0	0	0	9	9	0	9
h7	0	0	0	0	9	9	0	9
h8	0	0	0	0	9	9	0	9

C=0.915

(d)

Fig.8(cont'd)

5. DISCUSSION

The proposed matching technique assumes to memorize every conceivable projected shape of a polyhedron as reference data. (This might also happen in the human brain when a familiar object is memorized.) The advantage for this is that the technique is insensitive to the posture of a polyhedron. Even if it is upside-down, the matching proceeds without difficulty. As was shown in the experiment, the present technique is also applicable to occluded cases. All these come from that local constraints, namely adjacent relations of faces were taken into account to look for match. As Waltz[7] points out, the employment of local constraints is often effective or even powerful when eliminating unnecessary correspondences from a large number of possible correspondences. However, an appearance-oriented model offers a large number of data to be matched, which causes much run time. To overcome this, parallel architecture should be considered as Ullmann *et al.*[10] propose.

The experimental results were satisfactory. Extension of the present technique to curved objects need to be investigated.

6. CONCLUSION

A pattern recognition technique was proposed employing network constraint analysis. The technique was examined experimentally and satisfactory results were obtained. The proposed technique needs to be extended further to the curved object case.

REFERENCES

[1] Horn, B.K.P.: *Robot Vision,* The MIT Press, Mass. (1986).

[2] Pugh, A.(Ed.): *Robot Vision,* Springer-Verlag, New York (1983).

[3] Minsky, M.: "A framework for representing knowledge", *The Psychology of Computer Vision* (P. H. Winston, Ed.), McGraw-Hill, New York, 211-277(1975).

[4] Koenderik, J.J., Van Doorn, A.J.: "The internal representation of solid shape with respect to vision", *Biol. Cybernetics*, 32, 211-216(1979).

[5] Fekete,G., Davis,L.S.:"Property spheres: A new representation for 3-D object recognition", Proc. Workshop on Comput. Vision, 192-201(1984).

[6] Cherry, C., Vaswani, P.K.T. : "A new type of computer for problems in propositional logic, with greatly reduced scanning procedures", *Inform. Control*, **4**, 155-168(1961).

[7] Waltz, D.L. : "Understanding line drawings of scenes with shadows", *The Psychology of Computer Vision* (P. H. Winston, Ed.), McGraw-Hill, New York, 19-91(1975).

[8] Rosenfeld, A., Hummel, R. A., Zucker, S. W. : "Scene labeling by relaxation operations", *IEEE Trans.Syst.Man,Cybern.*,**SMC-6**,6,420-433(1976).

[9] Haralick, R. M., Shapiro, L. G. :"The consistent labeling problem: Part 1", *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-1**, 2, 173-184(1979).

[10] Ullmann, J. R., Haralick, R. M., Shapiro, L. G. : "Computer architecture for solving consistent labelling problems", *Comput. J.*, **28**, 2, 105-111 (1985).