Cooperative Recognition using Multi-View Images

Toshiyuki Kojoh*, Tadashi Nagata** and Hong-Bin Zha
*Dept. of Electrical Eng. Kyushu University
6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812, JAPAN

Abstract

We represent a method of 3-D object recognition using multi images in this paper. The recognition process is executed as follows. Object models as prior knowledgement are generated and stored on a computer. To extract features of a recognized object, three CCD cameras are set at vertices of a regular triangle and take images of an object to be recognized. By comparing extracted features with generated models, the object is recognized.

In general, it is difficult to recognize 3-D objects because there are the following problems such as how to make the correspondence to both stereo images, generate and store an object model according to a recognition process, and effectively collate information gotten from input images.

We resolve these problems using the method that the collation on the basis of features independent on the viewpoint, the generation of object models as enumerating some candidate models in an early recognition level, the execution a tight cooperative process among results gained by analyzing each image.

We have made experiments based on real images in which polyhedral objects are used as objects to be recognized. Some of results reveal the usefulness of the proposed method.

1 Introduction

In the research of computer vision, multiple images are frequently used for acquiring information that can't be obtained by using a single image. There are two approaches to multi image processing. The first is to measure range data or velocity vectors by means of stereopsis or optical flow computation. Ito *et al.* calculated range data on basis of a triangulation and edge images[1].Horn measured optical flow from a gradient of each image element intensity and a intensity change between images[2]. The second is focused on recognizing complex scenes by extracting image features with image processors, which respectively associate each of the images taken from different viewpoints and fuse the results. Chiou *et al.* proposed an approach for single polyhedron using three view analysis[3]. Mohan *et al.* presented a method for recognizing a 3-D object from 2-D images by index structural invariants from multiple views[4].

Our approach is the latter. In this approach, because they get various information concerned with an object to be recognized, and hence, it has been expected that the approaches would head for much better performances than single image methods. But, the expectaion has been disappointed because complexity of processing. We think that we should introduce a cooperative way such as images communicate each other and recognition results are fused in each process. However, due to the lack of systematic methods of cooperations among processors with different functions, the approach remains very different but the problem should be resolve in computer vision research.

We realized an effective recognition method according to following lines.

- Constructing an object to be recognized by calculating a two-face angle and a two-segment angle
- To reduce the number of matching operations in an image and the model correspondence, these surface and edge feature are grouped according to their similarity.
- A local process and a cooperative one are combined in the recognition.

In next section, we describes about an overview of general processing.

2 Overview

Fig.1 shows a processing overview. Main processing parts are a model generation, a feature extraction, and a recognition processing.

The model generation part gives a computer an object model as a prior knowledgement. It makes two type models concerning one object. One is described on the basis of its geometrical and structural properties and represented with a graph showing the surface and edge level features of the object. We call it a graph model. Another consisits of a two-face angle and a two-segment angle with a similar value and represented in the form of frequency distribution. We call it a distribution model.

The feature extraction part reconstructs a shape of an object reflected on an image. Extracted features are a two-face angle and a two-segment angle. Two-face angle is calculated from a gradient of a face and its adjacent relation with other faces. Two-segment angle is calculated from two edge parameters with the same endpoint and face.

In the recognition part, features extracted from the three different images are matched with the graph style model description using grouped models according to the similarity of the model features. Results of the matching processes are integrated at various levels, and the recognition procedure is immediately terminated when the integrated results are enough to prove the occurrence of the object to be recognized.

In the following sections, we describes about a detailed algorithm of each processing part.

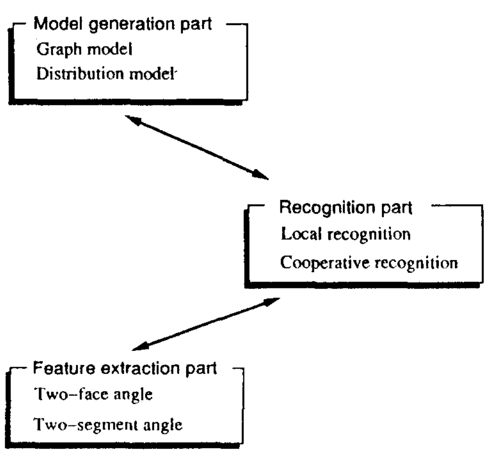


Fig. 1: An overview of a processing part

3 Model generation

3.1 Graph model

To recognize an object, it is necessary to store a knowledgement for its object and it is important how to represent an object model. As recognition process is executed on the basis of matching features such as two-face angle and two-segment angle, we must introduce a method which can represent them. So we try to represent a shape of object using a graph.

A graph model of a pyramid is illustrated in Fig.2. It has two levels, a face and an edge segment. The face level is composed of grouping segments belong to it. An object is constructed by grouping faces. A relation between the face level and the segment level is hierarchal.

In the face level, a normal vector of the surface is made an entry in a node, an adjacent surface and angle between these surfaces are made an entry in an arc.

In the segment level, endpoints of the segment are made an entry in a node, a connecting segment and an angle between these segments are entry in an arc.

It can be refered to data between two type levels using a pointer.

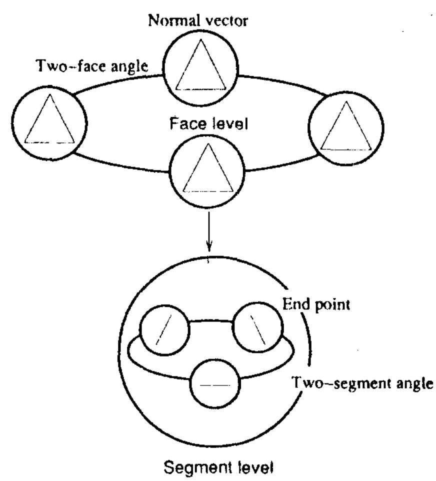


Fig. 2: A graph model of a pyramid

3.2 Distribution model

We generate another object model to reduce the number of matching operations. We call it a distribution model. This model is represented in the form of a frequency distribution table which is grouped with a similar feature to a two-face angle and a two-segment angle. This model is shown in Fig.3. A horizontal axis represents the two-face angle in the figure. A vertical axis represents the number of an existing angle. In the case of a segment level, A horizontal axis represents the two-segment angle in the figure, and a vertical axis represents the same as a surface level. As the matching operation is started from a part of a model with a similar feature to an extracted feature in an input image, an effective recognition can be realized. This effect is caused by using above mentioned models.

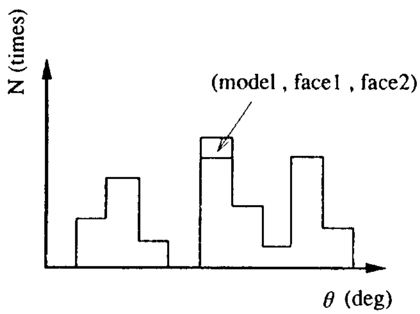


Fig. 3: A distribution model

4 Feature extraction

4.1 Extraction of a face information

We use a photometric stereo method[5] to find a gradient surface. It is measured resolving following expressions using each intensity value and light position parameters when an object is irradiated from three different positions.

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \frac{\rho}{\sqrt{p^2+q^2+1}} \begin{bmatrix} p_{s1}/s_1 & q_{s1}/s_1 & 1/s_1 \\ p_{s2}/s_2 & q_{s2}/s_2 & 1/s_2 \\ p_{s3}/s_3 & q_{s3}/s_3 & 1/s_3 \end{bmatrix} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \quad (1)$$

here,

$$s_i = \sqrt{p_{si}{}^2 + q_{si}{}^2 + 1} \qquad (i = 1, 2, 3) \quad (2)$$

I is a intencity value of an image element , ρ is a reflectance ratio, and p, q are gradient space parameters.

After a surface gradient on a image element is measured, we execute a clustering of an equivalent gradient value.

4.2 Extraction of a segment information

To extract edges of an object to be recognized, we use a filtering and Hough transformation[6] and calculate the following parameters.

$$\rho = x_i \cos\theta + y_i \sin\theta \quad (3)$$

$$\theta = \tan^{-1}\frac{\Delta y}{\Delta x} \quad (4)$$

Variables are shown in Fig.4. And endpoints of an edge is calculated using an image coordinate and their parameters.

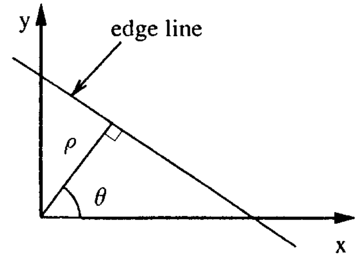


Fig. 4: Parameters of Hough transformation

4.3 Calculating of two type angles

In this subsection, we describe about calculating two types of angle. These are two-face angle and two-segment angle. Here, we consider two vectors, $\vec{a}$ and $\vec{b}$. In the case of a face level, the vectors are a normal vector of an adjacent surface. In the case of a segment level, the vectors correspond to directions of segments which belong to the same surface and have the same connecting point. An angle of these is expressed in (5).

$$\theta_f = \cos^{-1}\frac{\vec{a}\cdot\vec{b}}{|\vec{a}||\vec{b}|} + \frac{n\pi}{2} \quad (5)$$

A variable n is determined by considering an adjacent relation between a surface and a segment. As a result of calculating, a reflected object in each image is reconstructed in the same form as a graph model and a distribution model.

5 Recognition method

5.1 Flow of a recognition process

An object is recognized by the collation between models and features mentioned in section 3 and 4. Fig.5 shows a flow of a recognition method. It is mainly constructed by two level processes such as a local recognition and a cooperative recognition.

The former enumerates candidate models corresponding to an object reflected on an image in a single image matching. This level can't recognize its object because of an occlusion of itself and the viewpoint, but this method has a good algorithm that candidate models in all of a potential model to be matched is selected effectively.

The latter is a cooperative recognition. It eliminates mismatch models by getting and fusing informations from other images. As the result of their processes, an object is recognized.

In the following subsection, we explains an algorithm of two types of a potential recognition processing.

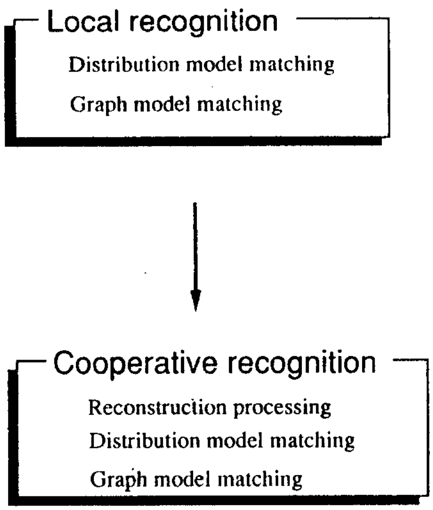


Fig. 5: A recognition process

5.2 Local recognition

A local recognition searches models corresponding to an object to be recognized in each image.

At first, from patterns reconstructed using a method mentioned in section 4, we generate their frequency distribution such as a distribution model. And the number of feature having the fewest value of frequency is searches in a distribution model. We call its process a distribution matching.

Second, if it is less than a threshold, feature patterns around its feature is compared with a graph model corresponding to an enumerated model. We call its process a graph model matching.

In the case that a visible face is a single or the number of a frequency in a distribution model is not less than a threshold, these processings don't execute.

5.3 Cooperative recognition

After a local recognition has been executed, an object viewed in an image is recognized in this recognition lebel.

Reconstruction processing reconstructs an occulusion part of its object using viewpoint parameters and surface and segment parameters. When a new two-face or two-segment angle is calculated, the distribution matching is executed.

A graph model matching is executed by the process which exchange reconstructed patterns and candidate models enumerated in a local recognition.

This processing is executed for two or three images.

6 Experiment and result

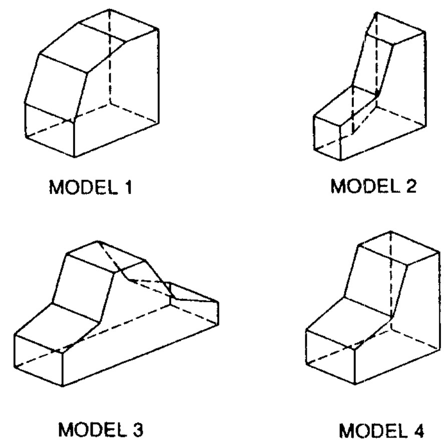
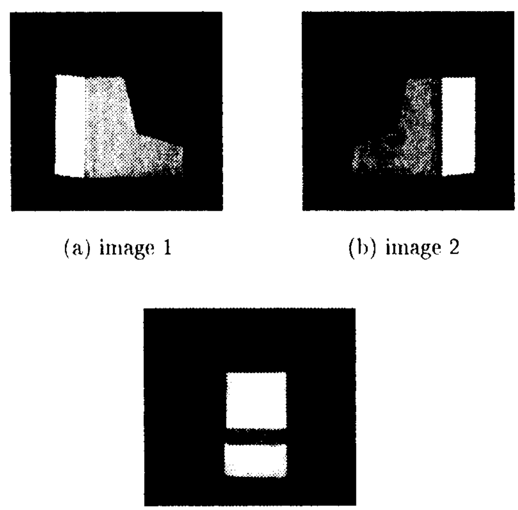


Fig. 6: Object models



(a) image 1

(b) image 2

(c) image 3

Fig. 7: Input images

We show that these proposed algorithms are more effective in the 3-D object recognition in this section. Object models considered for this experiment are illustrated in Fig.6. As these models have similar shapes, it is impossible to recognize an object by using only a single image, and a cooperative processing among images is necessary.

One of these is selected as an object to be recognized. Fig.7 shows input images from each viewpoint. CCD cameras are set at vertices of a regular triangle and the position of its object is the center of gravity. You may easily expect that the recognition is impossible by only a single image because many different models are considered as candidacy models. Model group enumerated as the candidate is illustrated in Fig.8. As a result of a local recognition processing, note that MODEL1 is enumerated as a candidate on account of an ambiguous processing because it can't be decided that a shape of an adjacent connected surfaces is a convex or a concave surface by only this processing. Executing of a cooperative recognition can eliminate ambiguous or incorrect processing results, and recognize the object exactly. Fig.9 illustrates a shape of an object reconstructed by a construction processing. Fig.10 illustrates models enumerated as results of a cooperative recognition with adjacent images. In this processing level, you may note the correspondence model narrows down the object. As a result of a cooperative recognition among three images, it is made clear that an object reflected an image is MODEL4.

We have made an experiment to show the effectiveness of the proposed cooperative recognition algorithm. Generally, the effectiveness can be shown by comparity matching process that we proposed with that of a method, in which matching is carried out between a graph model and feature patterns after the object to be recognized is reconstructed from three images. In the experiment, the number of matching process for the proposed method is 63, while that of the another method is 873.

7 Conclusions

In this paper, we described a method for recognizing 3-D objects by using multi-view image. it is made clear that this method can perform a correct and effective recognition by a tight cooperative processing among three images. In future, we will extend this method for more complicated scenes such as some 3-D objects overlapped.

References

[1] M.Ito *et al.* : "Three View Stereo Analysis", *Trans. PAMI* Vol.8 No4 pp.524-533 1986

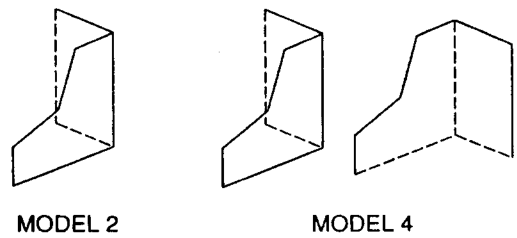
[2] B.K.T. Horn *et al.* : "Determining optical flow", *Artificial Intelligence* 17 pp.185-203 1981

[3] Rong-Nan Chiou *et al.* : "Polyhedron Recognition Using Three-view Analysis", *Pattern Recognition* Vol.25 No.1 pp.1-16 1992

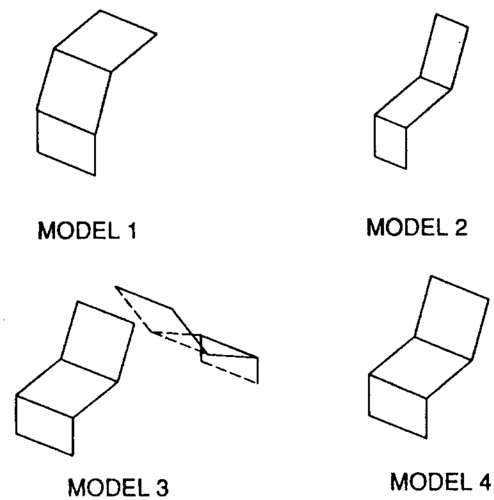
[4] R.Mohan *et al.* : "3D Object Recognition By Indexing Structural Invariants from Multiple Views", *In Proc.CVPR* pp.264-269 1993

[5] P.J.Woodham : "Analyzing Images of Curved Surface", *Artificial Intelligence* Vol.17 pp.117-140 1981

[6] R.O.Duda *et al.* : "Use of the Hough Transformation to Detect Line and Curves in Pictures", *Comm.ACM* ,Vol.15 No.1 pp.11-15 1972



(a) Image 1 , Image 2



(b) Image 3

Fig. 8: A result of a local recognition

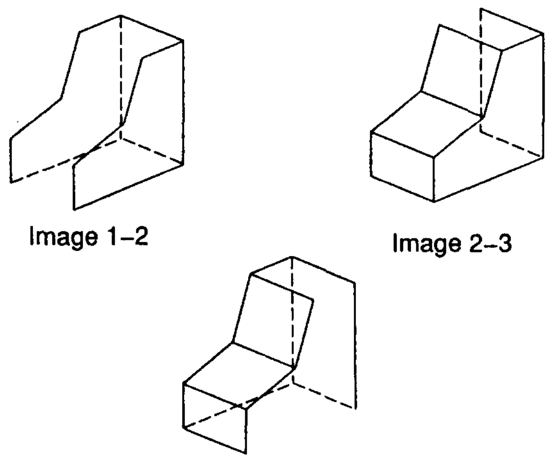
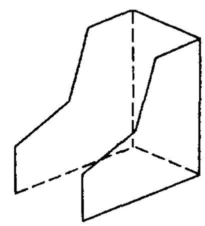
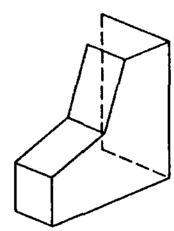


Fig. 9: A shape of an object by a reconstruction processing

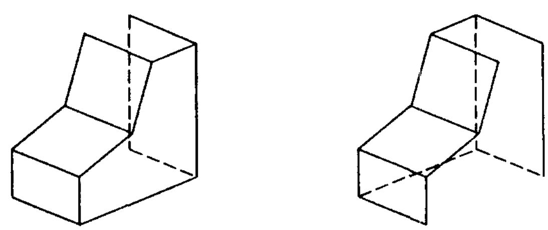


MODEL 4

(a) A candidate model of an image 1–2



MODEL 2



MODEL 4

(b) Candidate models of image 2–3 and 3–1

Fig. 10: A result of a cooperative recognition