N93-22194

3D Image Recognition Based on Fuzzy Neural Network Technology

Kaoru HIROTA, Kenichi YAMAUCHI, Jun MURAKAMI, Kei Tanaka
Dept. of Instrument & Control Engineering, College of Engineering
Hosei University
3-7-2 Kajino-cho, Koganei-city, Tokyo 184, Japan

ABSTRACT

3-D stereoscopic image recognition system based on fuzzy-neuralnetwork technology has been developed. The system consists of 3 parts; preprocessing part, feature extraction part, and matching part. Two CCD color camera images are fed to the pre-processing part, where several operations including RGB-HSV transformation are done. A multi-layer perceptron is used for the line detection in the feature extraction part. Then fuzzy matching technique is introduced in the matching part. The system is realized on SUN spark station and special image input hardware system. An experimental result on bottle images is also presented.

keywords: 3D image recognition, fuzzy matching, neural network

1. Introduction

The recent development of image processing and pattern recognition technology is remarkable. Many are put into the practical use in fields of industrial testing system, remote sensing and so on. It is difficult, however, to make a flexible vision system based on human experiences and human skilled knowledge. On the other hand, fuzzy logic and neural network technology are applied over a lot of fields including the control and the image recognition, where a human like processing is introduced. By combining the both techniques, 2-D image recognition system has been realized and reported [1].

In this paper a newly developed image recognition system using the technique of the binocular vision with fuzzy neuralnetwork methodology is presented. The system is realized on SUN spark station and special image input hardware system with 2 CCD color cameras. It consists of the following 3 parts; the pre-processing part where RGB-HSV transformation and other operations are done, the feature extraction part where a multi-layer perceptron is used for the line detection, and the matching part where a fuzzy matching algorithm is introduced. Finally several experimental results on bottle images are presented in order to confirm the availability of the proposed system.

2. Image recognition process

Fig.1 shows the outline of the presented image recognition process. It is roughly divided into 3 parts.

In the preprocessing part binocular images of each 512*512 pixels are taken by using 2 CCD color cameras. Several ordinary image processing operations and the concept of color fuzzy set are introduced in order to satisfy the quality requested in the feature extraction part. Then the contour features are extracted in the feature extraction part by using the multi layer type perceptron and the factorization technique based on fuzzy logic.

Finally a fuzzy matching algorithm is introduced in the matching part, where the result is presented in terms of fuzzy set.

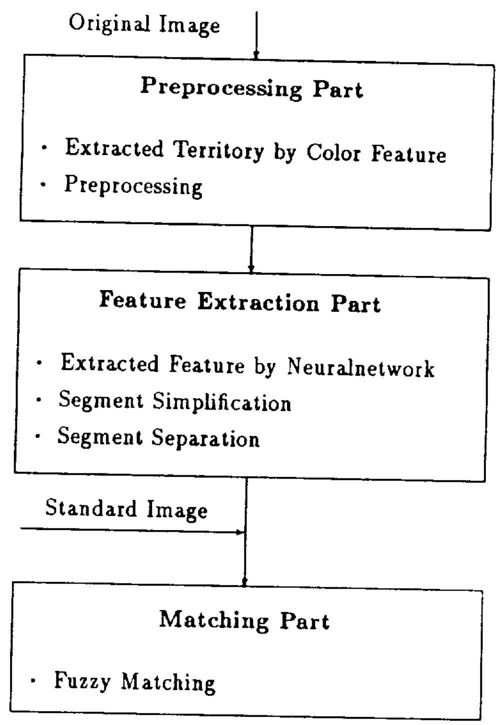


Fig.1 The outline of image recognition process

3.Colored region extraction based on color information

The input image from the camera is expressed by combining the RGB (Red, Green, and Blue) density. In the case of humanbeings color information is transmitted and is qualitatively recognized from eyes to a large brain (perception center). And a lot of models are proposed to explain the process. Here HSV (Hue, Saturation, and Value) hexagon cone color model [2] is used by introducing the RGB-HSV conversion. So the RGB color information is converted into three attributes of the hue, the saturation, and the value, where three attributes are defined by the membership functions which are shown in Fig.2.

When humanbeings extract the color features, the color distribution/tendency of the entire image is considered. For instance, when the image observed is composed of rather similar colors, then the color range to be recognized is set to be narrowed. Such characteristics are expressed by fuzzy rules. An example of fuzzy rules is shown below. By introducing the fuzzy matching technique, the feature colors are extracted.

[one example of fuzzy rules of feature color extraction]

IF the hue of the object is closely distributed

THEN the membership function of the hue should be narrowed.

4.Line segmentation using multi-layer perceptron

An image based on color information is converted to an line drawing image by using ordinary image processing technique. A multi-layer perceptron is applied to scanning the line drawing

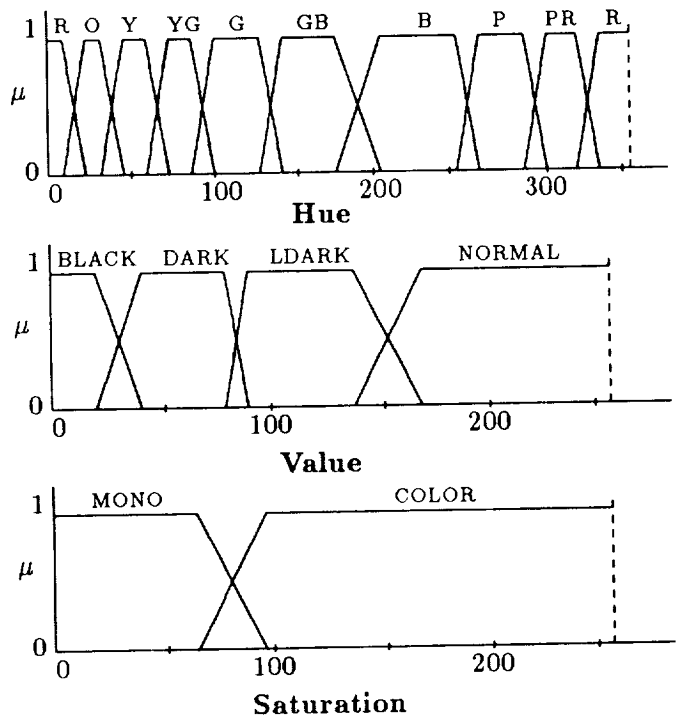


Fig.2 Membership functions of three color attributes

image and extract the directionality of lines. Here it should be noted that the non-learning data is generalized by the learning data in the Back Propagation Method of the perceptron model [3][4].

The input layer in the multi-layer perceptron corresponds to a part of the image. The teaching pattern in the learning process is a set of the typical lines of the input pattern, and the output pattern is the direction representing code called a chain code.

The outline of the multi-layer perceptron is summarized as follows: The output of the i-th neuron in the first layer is

$$u_i^l = f(s_i^l), \tag{1}$$

where

$$f(s) = \frac{1}{1+\exp(-s)}, \tag{2}$$

$$s_i^l = \sum_{j=1}^{N_{l-1}} w_{ij}^{ll-1} \cdot u_j^{l-1} - \theta_i^l . \tag{3}$$

The input of the (i,j) coordinate in the input frame is

$$u_i^1 = x(i\%W, i/W), \tag{4}$$

where % and / stand for the remainder and the quotient of division, respectively. The evaluation is given by

$$E = \frac{1}{2}\sum_{i=1}^{N_L}[u_i^L - y_i]^2, \tag{5}$$

where y_i stands for the value of the i-th neuron in the teaching pattern.

The input frame has a variable ratio which is determined by the ratio of the dark pixels to the bright pixels in the input frame. The position of the input frame is slightly changed in order to adjust the position of the center of gravity of the dark pixels to the middle of the input frame. The coordinate (G_x, G_y) of the center of the gravity of the dark pixels in the input frame, and the ratio S of the dark pixels to the bright pixels is calculated as follows;

$$G_x = \frac{\sum_{j=1}^{H}\sum_{i=1}^{W} x(i,j)\cdot i}{H\cdot \sum_{i=1}^{W} x(i,j)} \qquad (6)$$

$$G_y = \frac{\sum_{i=1}^{W}\sum_{j=1}^{H} x(i,j)\cdot j}{W\cdot \sum_{j=1}^{H} x(i,j)} \qquad (7)$$

$$S = \frac{\sum_{j=1}^{H}\sum_{i=1}^{W} x(i,j)}{H\cdot W} \qquad (8)$$

Based on these input values the output of the multi-layer perceptron is calculated. By moving the position of the input frame taking the output value into the consideration the line segment is traced. In the branching point of the line segment the input frame of the multi-layer perceptron also makes a branch and the line segment is traced in parallel.

5.The simplification of the extracted line segment data series

The output data series obtained by the multi-layer perceptron represents the directionality of the line segment. It is classified into a group of straight line, curved line, and corner by using the membership function shown in Fig.3.

The simplified data represent the geometrical feature of the input image. They can be understood in many ways with membership value.

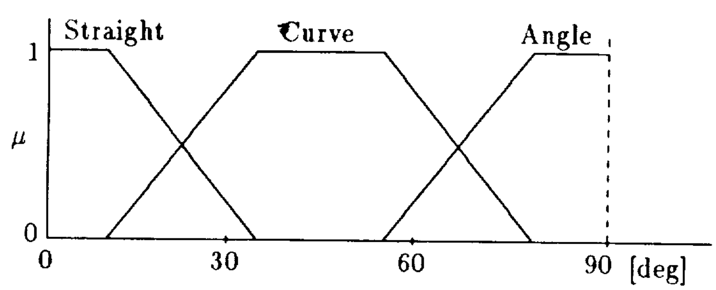


Fig.3 Membership function of geometrical features

6.Binocular stereoscopic vision

6.1 Correspondence between left and right images

Three dimensional binocular stereoscopic vision is realized by using in principle the difference between left and right images. But it is not so easy to make a correspondence of characteristic points between two images.

In this study the correspondence is made by using the simplified data series of both images mentioned in section 5. The line segment correspondence can be made based on the distance between line segments[5]. The both images are divided into several line segment blocks and the similarity between the blocks are calculated which generates sub blocks of line segments. Such a procedure continues and finally the correspondence of line segments are obtained, where fuzzy logic is introduced especially in the representation of the shape of line segment (c.f. Fig. 3) and correspondence operation.

6.2 Separation of objects

The distance between the camera and the object is calculated based on the information of line segment correspondence obtained in 6.1, where the method in projective geometry [6] are introduced. The position [a] in the 3D space and its corresponding position [b] in the image is connected by a translation matrix M

The matrix M can be calculated from the data of 6 points. Then the both images are transformed into 3D space by applying this translation matrix M. By doing a clustering procedure in 3D space the contour line of each object is extracted.

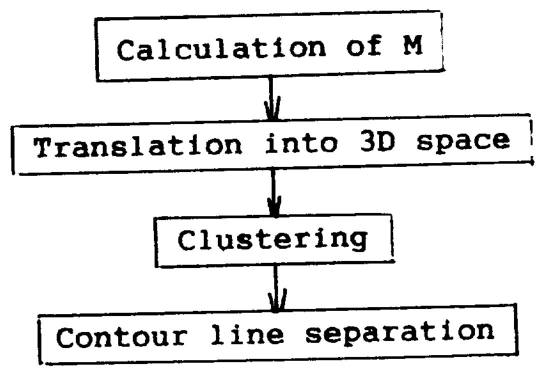


Fig.4 Contour line separation in 3D space

7. Fuzzy matching

Humanbeings can recognize the target object to some extent even if some part of it is hidden. Such functions are realized here by introducing fuzzy matching technique. The recognition result here is the similarity between the extracted information mentioned in the section 6 and the standard pattern information.

The standard pattern information consists of the type of line, the coordinates of starting point and end point, the length in the case of straight line, the curvature and the angle in the case of curved line and corner, and so on.

Firstly the segment with the minimum y coordinate is found and is checked if it is the top part of the object by observing the left and the right segments. Then the data series are divided into two parts. The similarity of the segment data against all standard pattern information is calculated. Then the relation between the segment data and the standard data with the maximum similarity is checked if there exist contradictions by considering other relations. By repeating this kind of procedure the final result is obtained. Table 1 shows the list of comparative features in the similarity calculation. Fig.5 shows their membership functions.

Table 1 A list of comparative features in the similarity calculation

Attribute	Comparative Feature	
Straight	· Start–End Coordinates	Δp
	· Length	Δl
	· Inclination	$\Delta\theta$
Curve	· Start–End Coordinates	Δp
	· Chord Length	Δl
	· Angle at the Circumference	$\Delta\theta$
Angle	· Angle	$\Delta\theta$

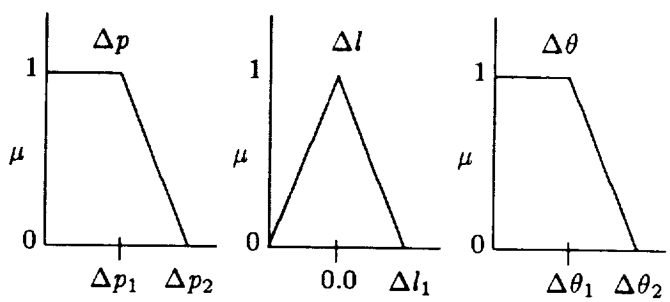


Fig.5 Membership functions of comparative features

8 Experimental result

In order to confirm the availability of this method several experiments have been done, among which one result using bottles is shown. Observed image of bottles consist of straight lines, curved lines with various curvature, and corners. There exist so many similarly looking different bottles. So they are good for testing the presented method.

Fig 6 shows several examples of original image observed by the CCD camera. Experiments were done for the single bottle, a pair of bottles, and three bottles. (The aim of latter two cases is to check the effect for occlusion.) The result is summarized in Table 2, which shows the validity of the proposed method.

9.Conclusion

Fundamental ideas and algorithms of 3D image recognition are proposed based on fuzzy neural network technique. A result of experiment on bottle images is also presented. The construction of real time 3D image recognition system for the purpose of robot vision is a part of future studies.

This study was performed through Special Coordination Funds of the Science and Technology Agency of the Japanese Government.

[References]

[1] K.Hirota, M.Katagai,K.Ikoma: 2D image pattern recognition based on fuzzy neuralnetwork technology, Proc. of 7-th Fuzzy Systems Symposium (SOFT), pp387/390 (1991) (in Japanese

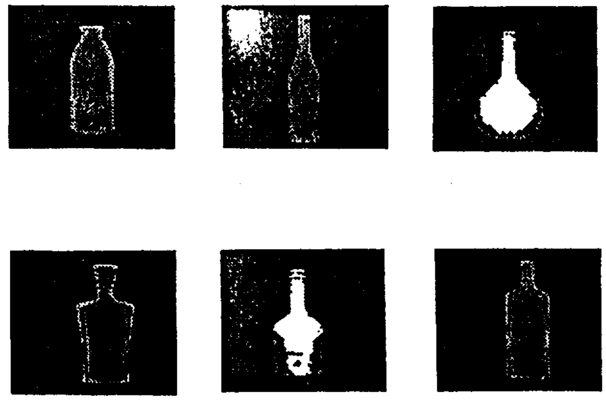


Fig.6 Sample patterns

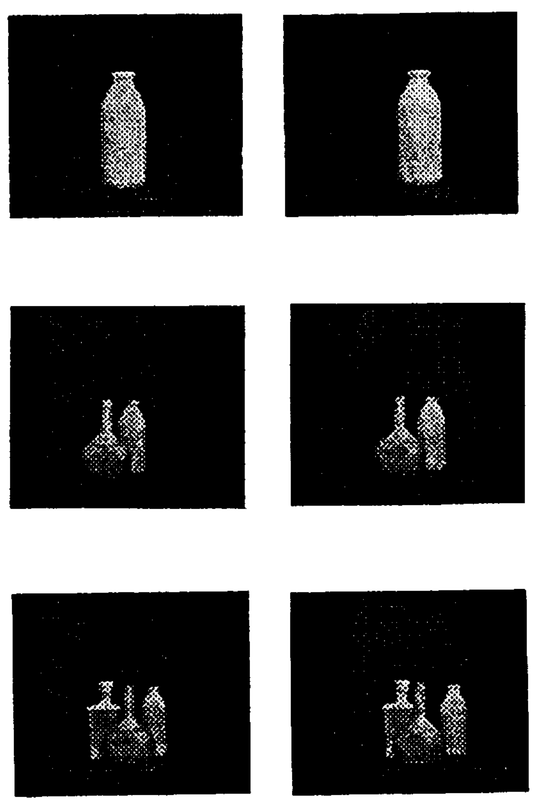
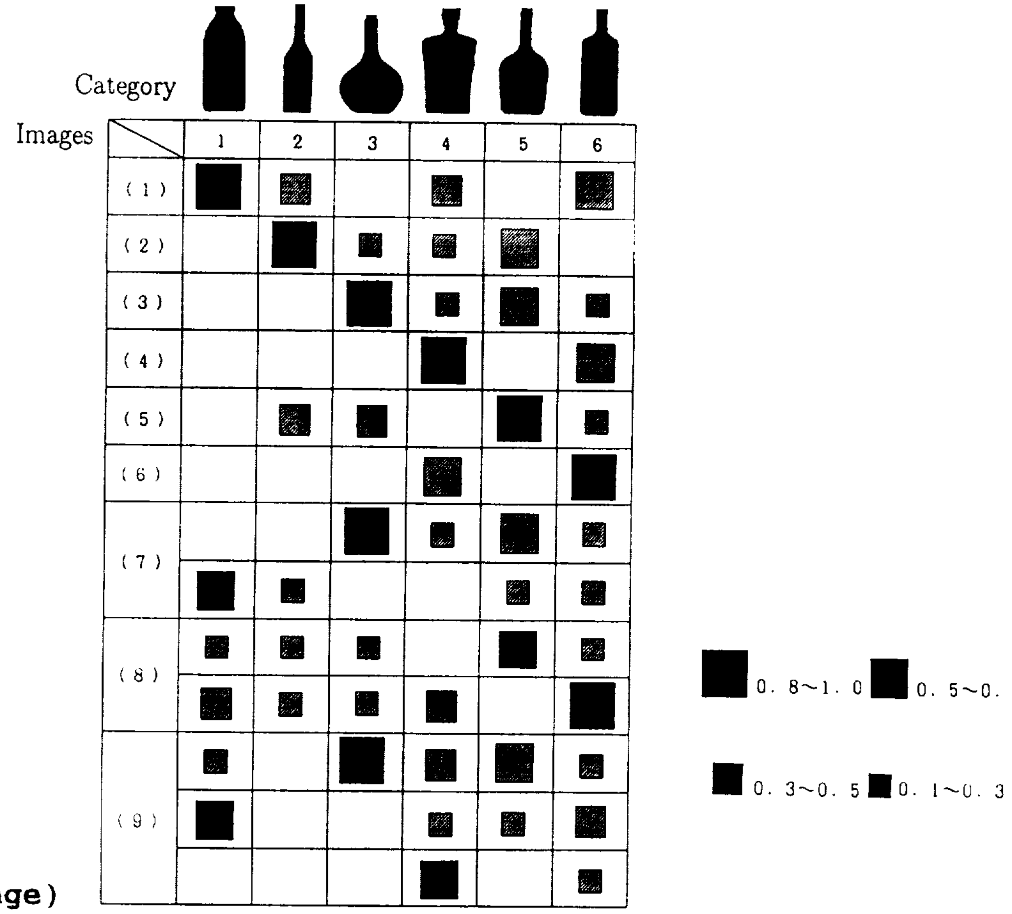


Fig.7 Examples of experimental data

Table 2 Recognition results



language)

[2] J.D.Foley,A.Vandam: Fundamentals of Interactive Computer Graphics, pp.613/620 (1983)

[3] F.Rosenblatt: The Perceptron, A Probabilistic model for information storage and organization in the brain, Physic. Rev. 65(6), pp.386/408 (1958)

[4] D.E.Rumelhalt: Learning representations by back-propagating errors, Nature 329-9, pp.533/536 (1986)

[5] H.Fukunaga,T.Kasai:Partition of line segment images and its application to stereo image processing, Tran.IECE (Japan)(D) J72-D2, pp866/872 (1989) (in Japanese language)

[6] K.Deguchi:Computer Vision, Projective Geometry for Graphics [IV], J.of SICE (Japan),30-3, pp.241/246 (1991) (in Japanese language)