A TWO LEVEL MODEL-BASED APPROACH OF OBJECT RECOGNITION FOR ROBOTICS APPLICATION

DING-CHUNG LIU
Chung Shan Inst. of Science and Tech.
P.O. Box No. 1
Lung-Tan, Taiwan, R.O.C.

BEHROUZ PEIKARI
Electrical Engineering Department
Southern Methodist University
Dallas, Texas 75275

Abstract

A two-level model-based approach for object recognition is presented. The central idea of the proposed approach is grouping of the models into subsets of in the first level and recognizing the shape and locating the position and orientation of the object in the second level. Simulation results presented illustrate the efficiency and accuracy of this algorithm.

1. Introduction

The model-based object-recognition system is one of the most promising approachs for robot vision applications. The block diagram of a typical model-based recognition system is shown in Figure 1. The key issue of the model-based approach is matching the unknown image with a set of predefined models of objects. The capability of recognizing objects accurately depends upon the adequacy of the model, the description of the image, and the matching algorithm used.

Simple objects can be recognized by using two-dimensional spatial representation of the model, while complex objects with occlusion in low contrast background need higher dimensional representation. The structural description of models in pattern recognition uses simple components to describe complex scenes or objects. These low-level components were first called atoms by T. Pavlidis in [1]. A related term is "primitive" which has been used in syntactic pattern recognition. In the proposed approach, both local features and relational features are used to construct the corresponding primitives. A crucial subsystem of the model-based system is the modeling of the shapes for the objects. Much of the recent work in this area is concentrated on developing structural models [2-6]. In the proposed method both numerical and syntactic descriptions are presented. The utilization of syntactic model is to guide the search of subgroups of models, and the numerical model is used to match the models in details.

The basic idea is to identify the structure of unknown object with a set of rules in the first level. In the second level, the proposed technique combines the constraint verification method and evaluating a criterion function to match the unknown object and the model for recognition as well as to locate the position of the object.

2. Model building and shape description

In the proposed approach, both numerical and syntactic representations are constructed to form the model. The model is constructed from the local features and will have the properties of rotation invariant, translation invariant, and scale invariant. The utilization of the length of the side and vertex angle of the polygonal shape as the local features are commonly used to achieve the rotation and translation invariant. For scale invariance, normalization is one of the methods to overcome this problem for closed shapes. In the application presented in this section, a triangle which is centered at each vertex point is selected as the local feature. Instead of using the side length of two segments that form the triangle, a ratio of these two segments is utilized to obtain the scale invariant property. After boundary points of the image are extracted, the proposed system performs linear polygonal approximation to find the vertice of the border. A sequence of vertex points with coordinates (x_1,y_1), (x_2,y_2), (x_3,y_3), ..., (x_m,y_m), will result, if there are m vertices in the shape.

The Euclidean metric distance between two points are used to compute the lengths of the side of the polygonal segments:

$$l_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (1)$$

The vertex angle is determined by two adjacent segments which form the angle at the vertices. More specifically, let

$$\alpha_i = atn2\left[\frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)}\right] \quad (2)$$

and

$$\beta_i = atn2\left[\frac{(y_i - y_{i-1})}{(x_i - x_{i-1})}\right] \quad (3)$$

The vertex angle thus can be computed from α_i

and β_i as:

$$\theta_i = \begin{cases} \alpha_i - \beta_i & \text{if } \alpha_i - \beta_i \geq 0 \\ 360^\circ + (\alpha_i - \beta_i) & \text{if } \alpha_i - \beta_i < 0 \end{cases} \quad (4)$$

In addition to the computation of side length and vertex angle, the ratio, R_i of two adjacent segments between the vertex angle is also computed.

$$R_i = \left[\frac{l_i}{l_{i-1}} \right] \quad (5)$$

The model in the system thus can be defined according to the numerical values which are computed from above equations. The numerical description model is defined as follows:

$$M_r = \left\{ N,S,P \right\} \quad (6)$$

where

N is the code of the name for the object,

S is the total number of the primitives of the shape, and

P is the description of each primitive which is in the form:

$$P = (i,x_i,y_i,\theta_i,R_i) .$$

In this 5-tuple, i is the series number of the primitive, x_i,y_i represent the x–y coordinate of the vertex point, θ_i is the vertex angle, and R_i is defined in Equation (5).

To deal with a large number of models in the system, the classification phase for object recognition is divided into a two level process, one is the organizing level and the other is the recognition level.

Organization and identification of models

The organization of models is conducted by investigating the existence of characteristic structures. Those characteristic structures are defined in [7]. The subset of models is then organized by examining these characteristic structures. The larger the number of characteristic structures defined, the more categories can be obtained. When the unknown object is identified as belonging to one of the possible subsets, then the models in that subset will be used in the next stage to classify the unknown object as one of these models.

When the unknown shape is analyzed and the possible subset of models are assigned, the problem of identification of the unknown shape is then transformed to a problem of "matching" the description of the unknown shape with those of the models in the assigned subset. In this paper, an algorithm is proposed to find the quality of correspondence that simplifies Bhanu's [8] method which uses constraint verification procedure for evaluating a cost function. Two primitives are called "related" if their distance is less than a given threshold δ_T.

$$d[P_i , Q_j] < \delta_T \quad (7)$$

The Algorithm

The algorithm (Fig. 2) starts from the unknown shape with the numerical description Q_j trying to match the first model P_i of the assigned subset. The first step of the algorithm is to check $d[P_i , Q_j]$ between P_i, Q_j, see if they are related. Start from i=1 and j=1, if P_1 and Q_1 are not related, then increase the attribute j with 1 and check the relation for the next pair of P and Q. If they are related, then go to step 2. In step 2, the quality of correspondence of P_i and Q_j will be examined. This quality of correspondence which was first defined by Bhanu and Faugeras [8] combines the quantity of the distance measure of local features θ_i and R_i with the association features θ_{i-2}, θ_{i-1}, θ_{i+1}, θ_{i+2} and R_{i-2}, R_{i-1}, R_{i+1}, R_{i+2}.

After obtaining the quality of correspondence for all P_i, $i = 1,2,\ldots\ldots,n$, the total distance between the unknown shape and the rth model can be computed based on the quality of correspondence.

Using the similarity measure, the identifying task can be achieved by assigning one of the models in the subset as the unknown shape.

When the type of the unknown shape is assigned from the above criterion, the algorithm for identifying the unknown shape from the models is completed. The next important task of the system is to locate this recognized shape in the image plane.

Position determination is an important task in robot vision applications. The information about the position of the identified object can be incorporated to perform the task of picking up the object in the workspace. Generally, position determination can be viewed as a "registration" of a sensed image of unknown shape to a stored model with reference data, or in other words, it is a problem of finding the scale, translation and rotation factors between the transformation of sensed shape and stored model. In this paper, the proposed algorithm uses several points with high correspondence value to compute the transformations and take the average of all the transformations for the final result. A detailed description of this position and orientation determination is given in [7].

4. Simulation results and analysis

In order to investigate the performance and the ability of the proposed algorithm, a set of 16 different images of objects are obtained and used to conduct the experiments. These objects consist of a set of tools and character shapes. In the training phase, the models of these 16 objects are constructed. The images of the objects are taken by the RTIPS (Real Time Image Processing System) in the image laboratory of the EE department, at SMU. The boundaries of these images are shown in Figure 3. The image is 256 by 256 pixels in size and with 256 gray levels. The RTIPS takes the image and converts it to digital form and transfers it to Micro-Vax computer for further processing.

The models then are constructed from boundary detection, linear polygonal approximation, and numerical and syntactic description steps. The output of boundary detection steps are the border pixels of the image with the coordinates with respect to each boundary pixel. The number of boundary points for these sets of sample images range from 447 pixels to 1014 pixels. The output of linear polygonal approximation are the coordinates of the polygonal vertices. The number of vertices for all 16 images varies from 5 to 30. Table 1 shows the number of boundary points and the number of vertices for all 16 models. The test images are taken from the same set of objects. For each object, 20 images are obtained in 20 different positions and orientations. These images are then processed with the boundary detection, linear polygonal approximation, numerical and syntactic descriptions, structural detection, and then matched with the assigned subset of models to identify the type of the object, and determined the position of each image with respect to the model.

The above procedures were implemented on Micro-Vax computer with Fortran 77 languages. The timing analysis of the program running on Micro-Vax is shown in Table 2. Out of 320 test instance, 309 are correctly recognized. The total recognition rate for the test is 96.5. The error analysis of position determination is also obtained by taking the average of each error computed from the comparison between the test image and model image. The mean rotation error is -2.5°, and the mean translation error are 3.21 pixels and -4.63 pixels in x and y axes respectively.

References

1. T. Pavlidis, "The Use of a Syntactic Shape Analyzer for Contour Matching," *IEEE Transaction Pattern Analysis and Machine Intelligence, pp. 307-310, Vol. PAMI-1, No. 3, July 1979.*
2. K.S. Fu, *Syntactic Pattern Recognition and Applications*, Prentice-Hall, New Jersey, 1982.
3. T. Pavlidis, "Algorithm for shape analysis of contour and waveforms," *IEEE Transaction Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 4, pp. 301-312, July 1980.
4. G. Medioni and R. Nevatia, "Matching Images Using Linear Features," *IEEE Transaction Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 675-685, November 1984.
5. T.W. Sze, and Y.H. Yang, "A Simple Contour Matching Algorithm," *IEEE Transaction Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 6, pp. 676-679, November 1981.
6. B. Peikari and F. Badi'i, "Functional Approximation of Planar Curves via Adaptive Segmentation," *Int. J. System Soc.*, vol. 13, No. 6, pp. 667-674, 1982.
7. D.C. Liu, *A Two Level Model-Based Approach for Object Recognition*, Ph.D. dissertation, E.E. Department, Southern Methodist University, Dallas, Texas, U.S.A., May 1988.
8. B. Bhanu and J.C. Ming, "Recognition of Occluded Objects: A Cluster Structure Paradigm," *Intl. Conf. on Robotics and Automation*, pp. 776-781, 1986.

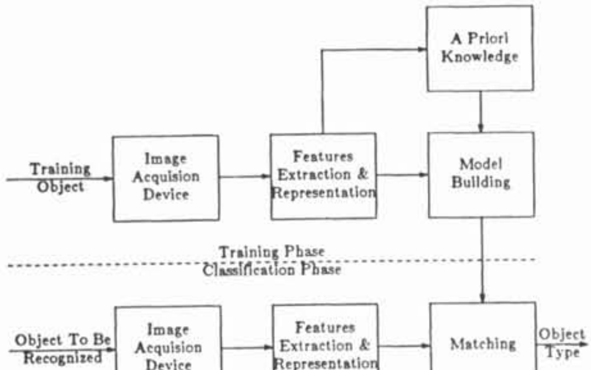

Figure 1 Typical block diagram for object-recognition system which utilizes model-based approach.

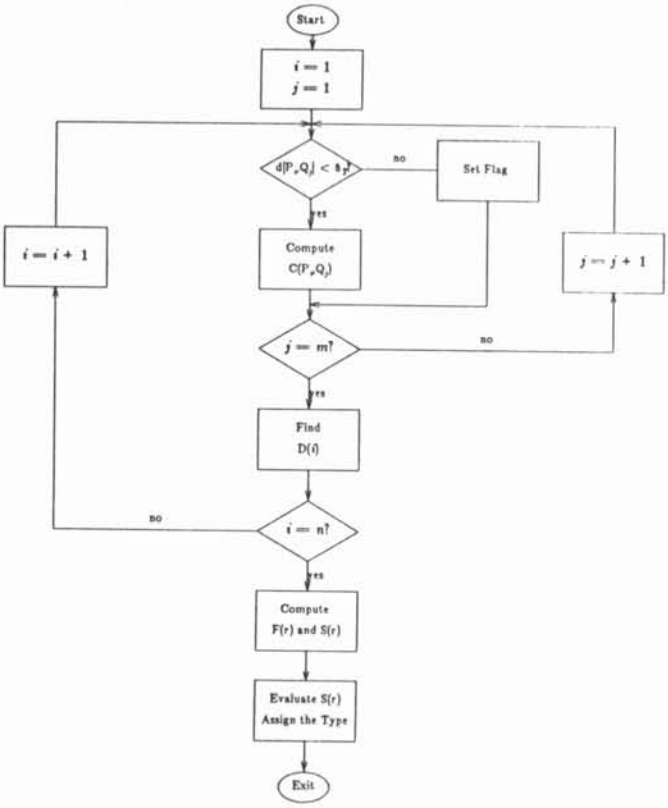

Figure 2 The block diagram of the algorithm to identify the unknown shape.

Table 1

The number of boundary points and vertices for 16 models

Type No.	No. of Boundary Points	No. of Vertices
1	506	5
2	1043	40
3	742	21
4	655	8
5	803	23
6	759	9
7	689	18
8	962	24
9	792	9
10	532	9
11	478	14
12	412	17
13	556	19
14	746	16
15	820	19
16	663	17

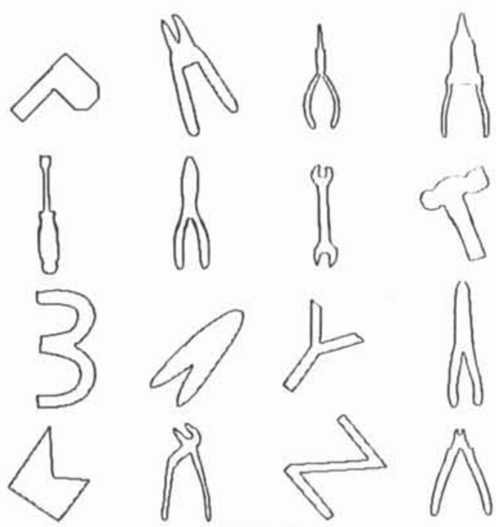

Figure 3
The border images for all 16 sample images

Table 2
Run time required in seconds when test on Micro-Vax

Name of Routine	Maximum time	Minimum time	Average
BDECT	1.3	1.1	1.2
PLYGS	0.7	0.4	0.65
NSREP	0.4	0.2	0.35
STRDET	0.7	0.4	0.6
MATCX	5.8	4.6	5.6
POSN	0.7	0.4	0.55
TOTAL TIME	9.6	7.1	8.95