IAPR Workshop on CV - Special Hardware and Industrial Applications OCT.12-14, 1988, Tokyo

FUZZY RECOGNITION OF 3_D OBJECTS

J.Mukherjee,P.P.Das and B.N.Chatterji

Dept. of Electronics and Electrical Communication Engg., I.I.T,Kharagpur-721-302,INDIA

I. Introduction

The essence of the 3-D object recognition technique is to derive an object representation from the sensor data and to match it with the stored models. There are different strategies, available in the literature, for object representation scheme and the matching technique. An object is represented by a set of primitives, such as points and lines [1], surfaces [2,3], voxels [4,5] etc. For wireframe representation an object is represented by a list of vertices and edges. Usually an object is recognised, in this case, by matching relational structures.

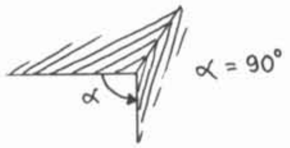
For 3-D object recognition, different matching strategies can be broadly categorised as (i) Graph matching strategy, (ii) Transformation of images into a feature space and thus match with the model feature vectors, & (iii) Knowledge based recognition strategy. Though syntactic technique is widely used for 2-D object recognition, its application to 3-D object recognition is limited due to its increasing complexity in handling 3-D object shapes.

In this paper we present a model-based recognition scheme for 3-D objects using fuzzy pattern matching techniques. The input to our vision system is given in the form of a 3-D binary array obtained from the range data of the 3-D object. Through a series of preprocessing, we convert this input into an equivalent 3-D wire-frame description, called the 3-D graph [6]. In this graph, the vertices represent the actual corners of the object and the edges are formed by connecting those vertices where a corresponding edge is present in the object itself. Since the problem is to identify a given object as one of the known object-models, we maintain an image library for the 3-D graphs of the known models. Hence the problem of recognition is reduced to fuzzy graph matching. These model graphs are characterised by feature vectors. Similarly, the derived 3-D graph can be characterised by a feature vector with a degree of fuzziness. And thus the matching can be done with a controllable degree of uncertainty. This scheme reduces the graph matching problem into a feature matching problem. Being a mathematical scheme it offers considerable computational advantages over the usual relational schemes.

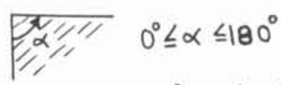
II. Representation of the objects in the world model:

Objects can be modelled in its wire frame structure with proper node characterizations. The six different node categories for this purpose are:

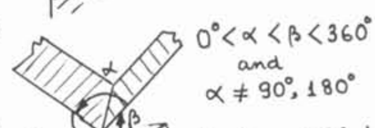
(i) C-Corner:



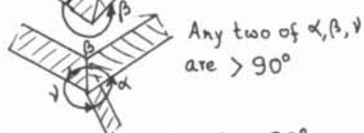
(ii) L-Corner:



(iii) V-Corner:



(iv) Y-Corner:



(v) X-Corner:

(vi) I-Corner:

Now we can characterise the object models by features formed by counting the number of corner or edge categories. We show an example of a 3-D 'B' in Fig. 1. Figure 1(a) shows the object, 1(b) shows the preprocessed 3-D graph and 1(c) its final form of wireframe structure. From 1(c) we have

$$FV_B = \overset{L\ \ C\ \ V\ \ Y\ \ X}{(0,\ 6,\ 4,\ 2,\ 0)}$$

$FM_B=$

	L	C	V	Y	X
L	0	0	0	0	0
C	0	7	4	2	0
V	0	4	2	4	0
Y	0	2	4	1	0
X	0	0	0	0	0

III. Fuzzy recognition scheme of an unknown object:

For an unknown object, we get its 3-D graph as G(V,E), where, V is the set of vertices and E is the set of edges and $v(i) \in V$, $(x_i, y_i, z_i) \in Z \times Z \times Z$.
Using the wire-frame structure, we compute the fuzzy membership $\mu(v(i),j)$ of its vertices $v(i), 1 \le i \le n$, to the corner type j, j = L, C, V, X, Y. This computation is based on the 2-D angle subtended by every pair of its edges meeting at v(i), where separate membership functions are used for acute, obtuse, right and straight angles. The method will be described later.

Then a working hypothesis Q is defined for the unknown object as,
$Q = \{(v(1),m(1)),(v(2),m(2)),\ldots,(v(n),m(n))\}$,
$m(i) \in \{L, C, V, X, Y\}, i = 1, 2, \ldots, n$.

The fuzzy membership $\mu(Q)$ of Q to the set of all possible hypotheses S is defined in terms of the $\mu(v(i),j)$'s as

$$\mu(Q) = \min_{i=1}^{n}\{\mu((v(i),m(i)))\}$$

Clearly there are 5^n different hypotheses for the object of which we are trying to select the "best" one. Taking a model O from the object-library we next compurte the similarity between the object-hypothesis Q and model O with respect to the feature vectors FV(Q) and FV(O) or feature matrices FM(Q) and FM(O). This similarity function will be denoted by $\mu(Q,O)$, mathematically it can be defined in various ways.

Two of such definitions which have been used according to the requirement are as follows:

$$(1)\ \mu(Q,O) = 1 - \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left\{1 - \frac{\min(Qn(i),On(i))}{\max(Qn(i),On(i))}\right\}^2} \quad ..(1)$$

$$(2)\ \mu(Q,O) = 1 - \frac{1}{t}\sqrt{\sum_{j=1}^{t}\max_{j\ \forall i}\left\{1 - \frac{\min(Qn(i),On(i))}{\max(Qn(i),On(i))}\right\}^2} \quad ..(2)$$

It can be verified that the above definitions are a normalised metric compatible to the membership $\mu(Q)$. Though it is defined for five (or fifteen) features, it can be immediately extended to more number of features for better reliability at the cost of extra computation.

Finally, the membership of the unknown model G to the selected model O is given by

$$\mu(G,O) = \max_{Q \in S}\{\min(\mu(Q),\mu_-(Q,O))\}$$

Since the size of S is exponential the computation of $\mu(G,O)$ is highly expensive. However, using a high threshold on the $\mu(v(i),m(i))$'s in the determination of $\mu(Q)$, we have been able to restrict the number of viable hypotheses to a reasonable limit without any noticeable degradation in the performance of the system. Once the $\mu(Q,O)$'s have been found we declare a set of close matches with G from our library accroding to high similarity values.

IV. Fuzzy Labelling of the corners:

From the preceeding section, it becomes clear that assigning a membership value of a corner to a particular type i.e., the computation of $\mu(v(i),j)$ is the prime factor for the computation of the membership of an unknown object to a model (i.e., $\mu(G,O)$). This membership value $\mu(v(i),j)$ depends upon the nature of the angle subtended by every pair of edges at that corner. We take four types of angles according to our requirement i.e., acute, right, obtuse and straight. We can have the following observations (in Table 1).

Table - 1

Corner Type	Fig.	No. of children	Angular requirement
L		2	(i) right = 1* or acute = 1 or obtuse = 1 (ii)straight = 0
C		>=3	(i) right >= 3
V		>=3	(i) right = 2 (ii) oblique = 1 or acute = 1 (iii)straight= 0
X		>=4	(i) straight = 2
Y		>=3	(i) oblique = 3 or (oblique = 2 and ((acute = 1) or (right = 1)))
I		<=3	(i) straight = 1

*here (right = 1) no. of right-angle = 1.

Hence the computation of membership value of a corner to be labelled with a particular corner type depends upon the computation of the membership value of the subtended angles at that corner to its type of angles. The membership function for an angle θ is defined from angular thresholds in terms of the cosine measure. For acute angle:

$$\mu_a(\theta) = 1 + \frac{\cos\theta - \cos(15)}{\cos(15) - \cos(0)}, \quad 0 \le \theta \le 15$$

$$= 1, \quad 15 \le \theta \le 75$$

$$= 1 + \frac{\cos\theta - \cos(75)}{\cos(75) - \cos(85)}, \quad 75 \le \theta \le 85$$

$$= 0, \quad \text{otherwise.}$$

The membership functions for right angle (μ_r), obtuse angle (μ_{ob}) & straight angle (μ_{st}) are defined analogously.

At any corner for each subtended angle (θ), the membership values corresponding to acute, right, obtuse and straight angles are computed. Then from the observations in Table 1, we can define the membership functions of any corner to a particular type as follow.

Let $\theta(j)$, $j=1,2,..,m$ be the subtended angle at any corner.

Let us define certain terms and notation as follows:

$$\text{Angle} = \{a, r, ot, st\}$$

$$\text{Max}(1,l) = \max_{j=1}^{m} \{\mu_l(\theta(j))\}$$

and any one of the $\theta(k)$'s for which $\mu_l(\theta(j))$ is maximum is removed to prevent its reconsideration for further calculation i.e., $\mu_l(\theta(k) = 0, \ \forall \ l \in \text{Angle}$.

$$\text{Max}(0,l) = \max_{\forall \in j} \{\mu_l(\theta(j))\}$$

Mnm(a,b) = minimum of a & b

Mxm(a,b) = maximum of a & b.

Mxm(a,b,c,d,...) & Mnm(a,b,c,d,...) are defined analogously.

d.f. = degree of freedom of the corner.

$$m = \binom{d.f.}{2}$$

Hence membership function of any corner to be assigned to C is:

C_mem = mnm(max(1,r),max(1,r),max(1,r)), if d.f.>=3

= 0 , otherwise.

Other membership functions are defined similarly.

Here corner type I denotes a fake corner & we have to exclude it during feature vector formation. The effect of exclusion of I in the computing feature matrix for a particular hypothesis, is rather complicated. We have to look for possible edges to be formed by deleting the corner labelled to I (an example shown in Fig. 2.).

If the sumtotal of all other corners are sufficiently more than I-type corners, this effect becomes less and the recognition system can tolerate its mere exclusion without readjusting the labelled edges & thus recomputing the feature matrix.

We will generate only those hypotheses Q, for which $\mu(v(i),j) >= \mu_{th}$ such that

$$\mu(Q) >= \mu_{th}.$$

Here μ_{th} is taken as 0.5.

V. In choosing the similarity function:

The similarity function must be so chosen that it can be compatible with $\mu(Q)$ and also it should give the true picture of the similarity between the temporarily unknown object representation (the hypotheses Q) and the objects in the world model (O).

For this we have defined two new normalised metric functions, as given in eqn. (1) & eqn. (2). While defining these metric functions we have to consider the relative occurrences of the corresponding components of the feature vectors and hence the absolute difference, between them cannot reflect their true similarity. That is why commonly used normalized Euclidean distance measure cannot be used in this regard. On the otherhand the applicability of the second measure is more appropriate in the edge-type feature model and the first measure to a vertex-type feature model. The reason behind it is that as the number of features increases, the non existance of any feature become common to both the feature vector and the similarity measure as defined by eqn(1) will always tend to higher values. But the second measure takes care of this.

Again, we may note that the non_existance of any feature vector, always contributes '1', which tends to higher membership values . Hence, before comparing between two feature vectors, both the feature vectors are translated by an unit vector i.e., for each component x_i.

$$x_i = x_i + 1;$$

VI. Results and discussions

In our vision system we have taken 3-D binary array of an object as our input to the vision system. These objects are transformed into a wire-frame structure and then matched with the stored object models. The similar preprocessing technique can also be implemented for range images. But for a range image the modelling requires the estimation of representative feature vectors from multiple viewing directions & storing them as a set of representative feature vectors & choosing the best match for the unknown object.

Here the results for the recognition of the objects represented on the form of 3-D binary array are shown in Table 2 and Table 3.

Table 2

	B.V*	S.V	cube.V	anvil.V
B	0.8002	0.5171	0.4753	0.3181
S	0.5603	0.7686	0.3968	0.4630
anvil	0.4073	0.4630	0.4205	1.0000
arrow	0.3866	0.4105	0.2650	0.6910

Table 3

	B.M*	S.M	cube.M	anvil.M
B	0.7388	0.4472	0.4300	0.4124
S	0.5000	0.6785	0.4385	0.4472
anvil	0.4472	0.5286	0.4385	1.0000
arrow	0.3997	0.5578	0.3895	0.5774

* _.V denotes model feature vector and _.M denotes model feature matrix.

Table 2 demonstrates the object matching using vertex type feature vector and the similarity measure defined in eqn. (1), while Table 3 demonstrates the object matching using edge-type feature matrix and the similarity measure defined in eqn (2). The relative comparisons from the tables show that with some threshold one can take decision for matching the object. As for best matching strategy, according to Table 2 and Table 3, B is best matched to B, S to S, anvil to anvil and arrow to anvil. For the best case anvil can be said partially matched to arrow (because they have very similar tail structures). This shows that our recognition strategy can be applied for occluded shape matching also, by selecting the suitable part of an image, which is the 'best' match for an object model and acknowledging its presence in the scene.

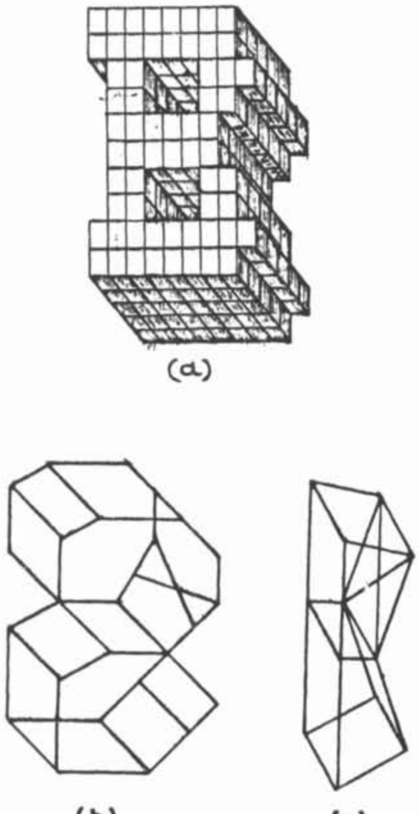


Fig. 1

```
             I
L|------|------|L      L|-----------|L
 |      |      |   ==>  |           |
L|------|------|L      L|-----------|L
        I          Fig.2
```

VII. References

[1] W.K.Gu and T.S.Huang, Connected Line Drawing Extraction from a Perspective View of a Polyhedron, in Proc. 1st. IEEE CAIA, 1984, 192_198.

[2] G.M.Radack and N.I.Badler, Local Matching of Surfaces Using Critical Points, Dept. Comp. Info. Sci., Univ. Penn., Tech. Rep., MS_CIS _ 84_13, 1984.

[3] F.Tomita and T.Kanade, A 3_D Vision System Generating and Matching Shape Descriptions in Range Images, in Proc. 1st. IEEE CAIA, 1984, 186_191.

[4] R.Nevatia and T.O.Binford, Description and Recognition of Curved Objects, Artificial Intelligence, vol. 8, 77_98, 1977.

[5] N.Badler, J.O.Rourke and H.Töltzis, A Spherical Representation of a Human Body for Visualising Movements, Proc. IEEE, vol. 67, 1397_1403, 1979.

[6] J.Mukherjee and B.N.Chatterji, A New Wire_Frame Object Modelling Technique, Proc. ISELDECS, IIT Kharagpur, Dec. 16_18, 1987, 375_377.