

Component Correspondence for an Imprecise Object Description and its Model

Naseem A. Khan

Department of Computer Science

Wayne State University

Detroit, MI 48202

and

Ramesh Jain

Electrical and Computer Engineering

The University of Michigan

Ann Arbor, MI 48109

March 1984

CENTER FOR ROBOTICS AND INTEGRATED MANUFACTURING

Robot Systems Division

COLLEGE OF ENGINEERING

THE UNIVERSITY OF MICHIGAN

ANN ARBOR, MICHIGAN 48109-1109

TABLE OF CONTENTS

1. INTRODUCTION	3
2. REPRESENTATION OF OBJECTS AND MODELS	4
3. BELIEF AND DISBELIEF VALUES	5
4. ESTABLISHING CORRESPONDENCE	8
5. SELECTION OF THE BEST MATCHED COMPONENTS	9
5.1. FORMULATION OF THE PROBLEM	10
5.2. HUNGARIAN METHOD TO SOLVE THE ASSIGNMENT PROBLEM	10
6. RESULTS	11
7. CONCLUSION	12

ABSTRACT

The component correspondence problem plays an important role in an approach for object recognition using models of objects. In this paper, we discuss our approach for solving the component correspondence problem. The features of regions in an image are assumed to be imprecisely recovered and are assigned certainty values in the range $[0,1]$. Each object has several components. The problem of establishing correspondence between the object components and the model component is solved using the *Hungarian method* which is commonly employed in solving assignment problems. The degree of match between a model and an object is obtained by combining the *belief value* and the *disbelief value* of individual components.

1. INTRODUCTION

Most approaches to the recognition of an object are based on the structure of the object and the characteristics of its components. Pattern recognition [DuH73] research tried to classify objects using some basic features. The limitations of statistical pattern recognition encouraged researchers to investigate approaches that tried to match components of an object with a model. The *degree of match* for an object and model was based on how well each individual component matched and on how closely the relations between components of the object matched with corresponding relations in the model. Many approaches were suggested that exploited this basic idea [DuH73,FiE73]. This idea can be conveniently represented by a graph theoretic framework. Each node of the *object graph* is a component, and the arcs between nodes represent the relations between the components. In the simplest, idealized situation, object recognition can be performed using graph isomorphism.

In realistic situations, one has to remember that the components and relations of the object may not match the model perfectly, some components may not even be present. Several approaches based on the stored description of objects have been suggested [BBF78, Bin82, Bro81, FiE73, Fis83, Gli83]. However, before we compute the degree of match between an object and a model, the *component-correspondence* problem must be solved. Both the object and the model have several components. Deciding which component of the model matches which component of the object, is the component-correspondence problem. This problem may prove to be difficult, particularly in images of a 3-dimensional scene in which objects may appear in random situations. Reddy and his co-researchers suggest that this problem should be solved using region features and search techniques [Pri76, Red78, RuR77].

It should be mentioned here that the *token-correspondence* problem, commonly faced by researchers in stereo and structure from motion, is quite different. In token-correspondence, the structure of the neighborhood of the points and knowledge about the imaging conditions play a dominant role. The component-correspondence addressed in this paper is at a symbolic level and exploits properties of regions. Another important difference is that the token-

correspondence is for two images, whereas the component-correspondence is for regions in an image and a model.

In complex situations, it helps to collect all supporting and opposing evidence for a hypothesis. The supporting evidence may be aggregated to obtain the *belief* in the hypothesis; and the opposing evidence may be aggregated to obtain the *disbelief*. The *belief* and *disbelief* represent the confidence that the hypothesis is true and false, respectively. The final confidence in the validity of the hypothesis can be obtained by combining them, as was done in [JaN79]. The idea of using both *belief* and *disbelief* values was successfully employed in MYCIN [Sho76]; Shafer has advocated the use of belief functions in the theory of evidence [Sha76].

We present a technique for matching imprecise object descriptions with models. The correspondence problem is solved using the *Hungarian* method. This method operates on certainty factor values obtained using *belief* and *disbelief* values for object-model component pairs to find the optimal model component to match an object component. The *belief* and *disbelief* values for an object-model component pair are obtained from the similarity of these components and their relations to other components. The geometric features of components are assigned values in the interval [0,1] by comparing them with ideal features; and are combined with other features using *fuzzy operators* [Kau75]. Our approach is strongly influenced by the *principle of least commitment* and employs an optimality criterion for establishing component correspondence.

2. REPRESENTATION OF OBJECTS AND MODELS

Suppose that regions in an image have been formed using a domain-independent method. Various properties of regions have been computed. The properties are represented in terms of the degree of the presence of a standard property. Thus a region is analyzed to find its *rectangularity*, *circularity*, and other similar properties. These properties are assigned values in the interval [0,1]. Some other numerical attributes may also be computed. The spatial relations between different regions of an object are found. This information is represented in the form of

Semantic Nets (SN), and is shown in Figure 1. The object description in the knowledge base are also stored in the form of a semantic net, as shown in Figure 2.

3. BELIEF AND DISBELIEF VALUES

For component matching, one has to consider the attributes of the components and the relationships of the components with other components. In the following, we present our approach for matching an input object description with a model in the KB. In Figure 3, we consider a simplified representation of an input object and the model. Each node in these graphs is labelled; it has some attributes attached to it. The branches are labelled by the relations that exist between different nodes. It is easy to see that in this representation each node is a region and a branch represents relationship between regions. Without any constraints, a node of the input graph could match any node of the model graph. Thus Figure 4 shows matching of input nodes with the model nodes. For recognition to occur, only one node in the input should match with one node in the model. This problem is solved at two levels: locally a pair should be compatible and globally the compatibility for the input and the model should be optimized. This can be achieved by assigning strengths to each pair based on the local properties and then using an optimization technique to select best matches. At the local level, we assign the confidence to a pair based on the evidence in support and against the match. The evidence supporting the match is aggregated to give the belief value, and the evidence against the match gives the disbelief value. This idea of combining supporting and opposing evidence was used in [JaN79].

If some of the features are missing in the input component, this is an *absence of information*. The features present in the input component contribute to the belief value. If the input component has some features, which are not present in the KB component, this contributes to the disbelief value.

Suppose, there are n components $C_{i1}, C_{i2}, \dots, C_{in}$ in the input, and m components $C_{K1}, C_{K2}, \dots, C_{Km}$ in knowledge base. Then BV_{ij} and DBV_{ij} are the belief value and disbelief value, obtained when component C_{Ki} in the KB is matched with C_{ij} in the input. For each

pair of components, three features are considered for matching. They are: 1) the geometrical shape, 2) the relative dimensions of the component, and 3) the spatial relationships of the component with other components. If the geometrical shapes do not match, then this indicates that the two components are entirely different. Only when the geometrical shapes match are any further attempts made to correlate two components through examination of the remaining features. A geometrical shape is associated with each component in the KB. If the input component does not include this geometrical shape in its IS-A list, then the KB component does not correspond to the input. Then a match between the component in the KB and that in the input is not feasible, and the belief value BV_{ij} and disbelief value DBV_{ij} for the pair are set to 0.

If the component in the input includes a geometrical shape similar to the component in the KB in its IS-A list, then the value of the instance-matching is given by

$$\text{INST-MATCH} = \text{feature certainty value associated with the IS-A}$$

In the next matching step, the relative dimensions of the component in the input are determined. If the component is a rectangle, ellipse, or a trapezoid, its length-to-width ratio is computed. If it is a circle, then the ratio of its diameter to the length of the largest, nearby rectangle, ellipse, or trapezoid is computed. The degree of dimension-matching is :

$$\text{DIM-MATCH} = 1,$$

if the relative dimension of the KB component matches that of the input component, *otherwise,*

$$\text{DIM-MATCH} = 0.5.$$

Because dimensions can be quite inexact, DIM-MATCH is assigned the temporizing value of 0.5 in the latter case. Because of this inexactness, we depend more on the shape of the component, and its spatial relationships to other components for matching. The final features that we attempt to match are the inter-relationships of components. For each component in the input, a list of relationships to other components is determined. A relationship k of the component in the KB being considered, is searched for in this list. If the relationship k is not found, there is an

absence of information. Then both the belief value R_k-BV , and disbelief value R_k-DBV , for the relationship k are assigned 0 values.

If the relationship is found, however, then the correspondence of components involved in this relationship is determined. If the relationship involves different component types in the two instances, there is refuting evidence, and we let

$$R_k-BV=0$$

$$R_k-DBV=INST-MATCH \cap DIM-MATCH$$

where \cap is the fuzzy *and* operator.

The disbelief value is always constrained by the match values for the shape and relative dimension features. To see this, let the feature value associated with a rectangular shape be 0.1, and let the relative dimensions match perfectly. The degree of certainty of the relationship then, must depend on the certainty that the component involved is actually a rectangle, and therefore, the disbelief value should not go above 0.1.

If the relationship involves similar components, then the degree of match of the component in the relationship, COMP-MATCH, is computed as follows:

$$COMP-MATCH=INST-MATCH \cap DIM-MATCH \cap FVrel$$

$$\cap FVinst-rel-comp \cap FVdim-rel-comp.$$

Here, $FVrel$ is the feature value associated with the relationship, $FVinst-rel-comp$ is the feature value associated with the component involved in the relationship, and $FVdim-rel-comp$ is the feature value associated with the relative dimension of the component in the relationship. If there are other components that have the same relationship with the KB component, then similar calculations are performed. Now we let

$$R_k BV = \frac{1}{K} \sum_{i=1}^K \underline{1} COMP-MATCH_i,$$

where K is the number of components that have this relationship. The average value is chosen since the relationship exists with several components and any one of these may be the correct instance. Also,

$$R_k - DBV = 0,$$

because there was no inexactness of information.

This process is repeated for all the relationships shown by the component in the KB. Finally, the belief value BV_{ij} and the disbelief value DBV_{ij} of the match between the i^{th} component in the input and the j^{th} component in the KB are computed. These are calculated as the average values of the belief and disbelief values for all the matching features. They are as follows:

$$BV_{ij} = \frac{(INST-MATCH + DIM-MATCH + REL_1-BV + REL_2-BV + \dots + REL_p-BV)}{(p+2)}$$

$$DBV_{ij} = \frac{(REL_1-DBV + REL_2-DBV + \dots + REL_p-DBV)}{p}$$

where p is the total number of relationships.

INST-MATCH and DIM-MATCH are not considered in the computation of DBV_{ij} because they do not contribute to the disbelief value.

4. ESTABLISHING CORRESPONDENCE

For each pair of components between the input and knowledge base, the corresponding Certainty Factor Value (CFV) is found using,

$$CFV_{ij} = BV_{ij} - DBV_{ij} \quad (i=1,2,\dots,m; j=1,2,\dots,n)$$

The degree of match between the i^{th} KB component and j^{th} input component is given by CFV_{ij} . The pairs with the best match can be determined from their CFVs using a global approach. What we need is an approach that selects those pairs that result in the maximization of the confidence for the input-model match. The problem of determining the best matching

pairs is formulated as an assignment problem, which is solved using the Hungarian Method, and is described in the next section.

After the optimization has been applied to give the pair (i,j) we can compute the overall belief and disbelief value for the object with respect to the model under consideration. Depending on the number of components in the input and the model, following cases arise:

CASE 1: $m=n$

$$FINAL-BV = \frac{1}{m} \sum_{i=1}^m BV_{ij}; \quad FINAL-DBV = \frac{1}{m} \sum_{i=1}^m DBV_{ij}$$

If there are additional components present in the input they contribute to the disbelief value of the object being considered. Thus, the following :

CASE 2: $m > n$

$$FINAL-BV = \frac{1}{m} \sum_{i=1}^m BV_{ij}; \quad FINAL-DBV = \frac{1}{m} \sum_{i=1}^m DBV_{ij}$$

CASE 3: $m < n$

$$FINAL-BV = \frac{1}{m} \sum_{i=1}^m BV_{ij}; \quad FINAL-DBV = \frac{1}{n} ((n-m) + \sum_{i=1}^m DBV_{ij})$$

5. SELECTION OF THE BEST MATCHED COMPONENTS

The degree of match between a KB component and an input component is the corresponding CFV. The best match between an input object and a KB object is obtained when the sum of the degrees of match between the corresponding components of the objects is the highest. This sum can contain no more than one occurrence of every input and every KB component, i.e. each component of the input object can be matched to at most one component of the KB object, and two different components from the input cannot match to the same component of the KB. There are many possible combinations to compute this sum, each way yielding a different sum and therefore a different degree of match. The highest value of the sum corresponds to the best match.

5.1. FORMULATION OF THE PROBLEM

Our objective in the object matching problem is to achieve an optimal match. The problem of selecting the best match for components of the input and KB object, can be solved using the techniques of linear programming. The problem is considered as an assignment problem, which is a special case of the transportation problem. The assignment problem is solved using the Hungarian Method.

5.2. HUNGARIAN METHOD TO SOLVE THE ASSIGNMENT PROBLEM

The objective in the assignment problem [CoS74] is very similar to the task of assignment of m jobs to m persons. The problem of matching the components in the input image and the knowledge base can be formulated as follows:

$x_{ij} = 1$ if a KB component i is paired with input component j

$= 0$ otherwise

C_{ij} = Degree of matching between the KB component i and the input component j

The maximization problem is converted to a minimization problem before applying the Hungarian Method. The costs C_{ij} in the minimization problem correspond to the *degree of disparity* between the components in the input and KB objects, which is $1 - CFV$. The minimization problem is shown below:

Minimize $z = \sum_{i=1}^m \sum_{j=1}^m C_{ij} x_{ij}$ subject to

$$\sum_{j=1}^m x_{ij} = 1 \quad (i=1,2,\dots,m)$$

$$\sum_{i=1}^m x_{ij} = 1 \quad (j=1,2,\dots,m)$$

$$x_{ij} \geq 0 \quad (i=1,2,\dots,m, j=1,2,\dots,m)$$

where C_{ij} is now the degree of disparity between the KB component i and the input component j .

The Hungarian Method [BrH77], developed for solving this assignment problem, has been found to be computationally very efficient. The algorithm in [BrH77] can handle cases of equal and unequal number of components. Due to space limitations, we will not give this algorithm here.

6. RESULTS

This matching is part of a object recognition system, discussed in [K84]. The system has been implemented using LISP. We performed several experiments using symbolic input. Our aim in these experiments was to study the efficacy of our approach for object recognition in presence of variability in the input data. We considered arbitrary values for the attributes of regions and different relationships between regions for a class of objects. The results of recognition are very encouraging.

We showed a sample input object description in Figure 1 and a model in the KB in Figure 2. By comparing different components using belief and disbelief values, as discussed in preceding sections, we obtain the following table:

	WB1	WB2	WB3
ONE	0.8	0.0	0.0
TWO	0.0	0.8	0.73
THREE	1.0	0.8	0.73

The above table is input to the Hungarian method. The result is the pairs (*WB1 ONE*), (*WB2 TWO*), (*WB3 THREE*) with their degrees of match 0.8, 0.8, 0.73, respectively.

It should be mentioned here that the object recognition is not achieved at this stage. Other knowledge sources are also used at this level and then the the final object recognition is performed using a two-level multiple knowledge sources based approach [K84].

7. CONCLUSION

In this paper, we present a method for matching an imprecise object description with models of known objects. Both the objects and their models are represented using semantic nets. First an object component is matched with each component of a model. The similarity and dissimilarity of the object-model component pair is represented using belief and disbelief values. After all such values have been obtained by comparing each object component with each model component, Hungarian method is used to establish object-model component correspondences. Final belief and disbelief values for the match of the object with the model are obtained by combining the belief and disbelief values for the corresponding components. The object recognition will take place after such belief and disbelief values have been obtained for each model in each knowledge source. The final recognition will be achieved using distributed problem solving, and is considered beyond the scope of this paper.

REFERENCES

[BBF78]

Ballard, D.H., Brown, C.M. and Feldman, J.A., "An approach to knowledge-directed image analysis", *Computer Vision Systems*, Academic Press, pp271-282, 1978.

[Bin82]

Binford, T.O., "Survey of Model-Based Image Analysis Systems", *Int. J. of Robotics Research*, Vol. 1, no. 1, pp18-64, 1982.

[BrH77]

Bradley, S.P. and Hax M., *Applied Mathematical Programming*, Addison-Wesley Publishing Co., 1977.

[Bro81]

Brooks, R.A., "Symbolic Reasoning Among 3D Models and 2D Images", Stanford AIM-343, 1981.

[CoS74]

Cooper and Steinberg, *Methods and Applications of Linear Programming*, W.B. Saunders Co., Philadelphia, PA, 1974.

[DuH73]

Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

[FiE73]

Fischler, M.A. and Elschlager, R.A., "The Representation and Matching of Pictorial Structures", *IEEE Transactions on Computers*, C-22, pp67-92, 1973.

[Fis83]

Fisher, R.B., "Using surfaces and object models to recognize partially obscured objects",
Proc. IJCAI-83, pp989-995, 1983.

[Gli83]

Glickman, J., "Using multiple information sources in a computational vision system",
Proc. IJCAI-83, pp1078-1080, 1983.

[HaR78]

Hanson, A.R. and Riseman, E.M., "VISIONS: A computer system for interpreting scenes",
Computer Vision Systems, Academic Press, 1978.

[JaN79]

Jain, R. and H.-H. Nagel, "On the Accumulative Difference Pictures for the Analysis of
Real World Scene Sequences," *IEEE Trans. on Pattern Analysis and Machine Intelligence*,
vol. PAMI-1, No. 2, pp. 206-214, 1979.

[Kau75]

Kaufmann A., *Introduction to the Theory of Fuzzy Subsets*, Vol. 1, Academic Press, 1975.

[K84]K (To maintain anonymity of authors, as suggested by AAAI, we are not giving complete

information at review time.) *Object Recognition Using Distributed Problem Solving, Ph.D.
Thesis, Department of Computer Science, Jan. 1984.*

[Pri76]

Price, K., "Change Detection and Analysis of Multi-Spectral Images", Ph.D. Thesis, Dept.
of Comp. Sci., Carnegie-Mellon University, Pittsburgh, PA, 1976.

[Red78]

Reddy, D.R., "Pragmatic aspects of machine vision", *Computer Vision Systems*, Academic

Press, pp89-100, 1978.

[RuR77]

Rubin, S.M. and Reddy, R., "The Locus Model of Search and its Use in Image Interpretation", Proc. of IJCAI-77, pp590-595, 1977.

[Sha76]

Shafer G., *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[Sho76]

Shortliffe, E.H., *MYCIN: Computer-Based Medical Consultations* American Elsevier, 1976.

UNIVERSITY OF MICHIGAN



3 9015 03023 8250