SHAPE MATCHING OF 2-D OBJECTS USING A HIERARCHICAL STOCHASTIC LABELING TECHNIQUE

Bir Bhanu

Digital Systems Department
Ford Aerospace & Communications Corporation
Newport Beach, CA 92660

Olivier D. Faugeras

INRIA, Domaine de Voluceau
B.P. 105, 78150 - Le Chesnay
France

ABSTRACT

A stochastic labeling technique to do shape matching of nonoccluded and occluded 2-D objects is presented. The technique explicitly maximizes a criterion function based on the ambiguity and inconsistency of classification. The technique is hierarchical and uses results obtained at low levels to speed up and improve the accuracy of results at higher levels. This basic technique has been extended to the situation when various objects partially occlude. In such a case several hierarchical processes are executed in parallel for every object participating in the occlusion and are coordinated in such a way that the same segment of the apparent object is not matched to the segments of different actual objects. Examples taken from aerial and industrial images are given.

I. INTRODUCTION

In this paper we address the "segment matching" problem of shape matching by using a hierarchical stochastic labeling technique. The class of shapes that we consider are represented by simple closed curves which are approximated by polygons. The shapes are allowed to have undergone translation, rotation, scale and significant changes in the shape.

II. SHAPE MATCHING OF NONOCCLUDED OBJECTS

We present a two stage hierarchical stochastic labeling method for matching the segments of a model against the segments of an observed object. Let $T = (T_1,\ldots,T_N)$ and $O = (O_1,\ldots,O_{L-1})$ be the polygonal path representation of the model and the object respectively, where T_i and O_j are line segments, $i = 1,\ldots,N$ and $j = 1,\ldots,L-1$. Model elements will be referred to as units and object elements as classes. We are trying to identify part of the model T within the observation O. We are therefore, trying to label each of the segments T_i ($i = 1,\ldots,N$) either as a segment O_j ($j = 1,\ldots,L-1$) or as not belonging to O (label O_L = Nil). Each segment T_i therefore has L possible labels. For every segment T_i we compute a set of L positive numbers $p_i(\ell)$, $\ell = 1,\ldots,L$ forming a vector $p_i = [p_i(1),\ldots,p_i(L)]^T$. $p_i(\ell)$ can be thought of as the probability of labeling the segment T_i as O_ℓ. The set of all vectors $\vec{p}_i$ ($i = 1,\ldots,N$) is called a stochastic labeling of the set of units.

Initially the stochastic labeling is ambiguous (except in some very special cases) and we make it evolve toward a less ambiguous labeling by comparing the local structure of T and O. From now on the indexes i are taken modulo N. To every segment T_i, we associate the two neighboring segments T_{i-1} and T_{i+1}. In order to compare the local structures of T and O we define two compatibility functions C_1 and C_2 of $S_2 \times O^2$ and $S_3 \times O^3$ into [0,1] where, S_2 and S_3 are two subsets of T^2 and T^3 defined by,

$S_2 = \{(T_i, T_j)\}$, $i=1,\ldots,N$, $j=i-1$ or $i+1$

$S_3 = \{(T_i, T_{i-1}, T_{i+1})\}$, $i=1,\ldots,N$

The compatibility function $C_1(T_i, O_k, T_j, O_\ell)$, $j=i-1$ or $i+1$ and $C_2(T_i, O_k, T_{i-1}, O_\ell, T_{i+1}, O_m)$ will be denoted more simply by $C_1(i,k,j,\ell)$ and $C_2(i,k,i-1,\ell,i+1,m)$. C_1 and C_2 take values between 0 and 1. $C_1(i,k,i-1,\ell)$ measures the resemblance of the set $\{T_i, T_{i-1}\}$ with the set $\{O_k, O_\ell\}$. A good (bad) match means that the value of C_1 is close to 1, (0). As described in [1,2] we can associate to every segment T_i a compatibility vector $\vec{q}_i = [q_i(1),\ldots,q_i(L)]^T$. Intuitively this vector represents what the neighbors of segment T_i (that is to say segments T_{i-1} and T_{i+1}) "think" about the way it should be labeled whereas $\vec{p}_i$ represents what the segment T_i "thinks" about its own labeling.

Mathematically speaking we compute

$$Q_{ij}(k) = \sum_{\ell=1}^{L} C_1(i,k,j,\ell)\,p_j(\ell), \quad j = i-1, i+1;\ i = 1,\ldots,N;\ k = 1,\ldots,L \tag{1}$$

$$Q_i^{(1)}(k) = \tfrac{1}{2}(Q_{i\,i-1}(k) + Q_{i\,i+1}(k)) \tag{2}$$

$$Q_i^{(2)}(k) = \sum_{\ell_1,\ell_2=1}^{L} C_2(i,k,i-1,\ell_1,i+1,\ell_2)\,p_{i-1}(\ell_1)\,p_{i+1}(\ell_2) \tag{3}$$

82CH1761-6/82/0000/0688$00.75 © 1982 IEEE

The numbers $Q_i^{(1)}(k)$ and $Q_i^{(2)}(k)$, $k = 1,\dots,L$ are positive. The idea is that they are large when the probabilities of the labels of the neighbors of T_i compatible with label O_k are large and small otherwise. The numbers $Q_i^{(1)}(k)$ and $Q_i^{(2)}(k)$ are normalized so that they add up to 1 yielding two vectors $\vec{q}_i^{(1)}$ and $\vec{q}_i^{(2)}$ such that,

$$q_i^{(j)}(k) = \frac{Q_i^{(j)}(k)}{\sum_{\ell=1}^{L} Q_i^{(j)}(\ell)}, \qquad j = 1,2 \quad k = 1,\dots,L \tag{4}$$

It is desired to decrease the discrepancy between what every segment T_i thinks about its own labeling ($\vec{p}_i$) and what its neighbors think about it ($\vec{q}_i^{(j)}, j = 1,2$). A good "local" measure of ambiguity and inconsistency is the inner product $\vec{p}_i \cdot \vec{q}_i^{(j)}$, $j = 1,2$. By computing the average over the set T of these local measures we obtain two global criteria:

$$J^{(j)} = \sum_{i=1}^{N} \vec{p}_i \cdot \vec{q}_i^{(j)}, \qquad j = 1,2 \tag{5}$$

The problem of labeling the segments T_i is therefore equivalent to an optimization problem: given an initial labeling $\vec{p}_i(0)$, $i = 1,\dots,N$, find a local maximum of the criteria $J^{(j)}$ ($j = 1,2$) closest to the original labeling $\vec{p}_i(0)$ subject to the constraints that p_i's are probability vectors. Since C_2 is a better measure than C_1 of the local match between T and O we are actually interested in finding local maximum of the criterion $J^{(2)}$. On the other hand maximizing $J^{(1)}$ is easier from the computational standpoint. We therefore use the following hierarchical approach: starting with an initial labeling $\vec{p}_i(0)$, we look for a local maximum $\vec{p}_i(1)$ of the criterion $J^{(1)}$. This labeling is less ambiguous than $\vec{p}_i(0)$ in the sense that many labels have been dropped (their probabilities $p_i(k)$ are equal to zero). We then use the labeling $\vec{p}_i(1)$ as an initial labeling to find a local maximum of the criterion $J^{(2)}$. The computational saving comes from the fact that the values C_2 corresponding to probabilities $p_{i-1}(\ell_1)$ or $p_{i+1}(\ell_2)$ equal to zero are not computed. The problem of maximizing (5) is efficiently solved using the gradient projection method [1,2].

Initial probabilities are obtained by computing the error between the feature values of the model and the object. Features used are length of a segment, intervertices distance, slope of a segment, and angles between segments called the interior angle and exangle [2]. Computation of compatibilities C_1 and C_2 involves binary and ternary relations respectively. There are several methods of computing them which are based upon finding the transformations and computing the errrors [2]. The problem of assigning initial probability and compatibility to the nil class is solved in [2].

III. SHAPE MATCHING OF OCCLUDED OBJECTS

Occlusion problem requires that the following two conditions be satisfied.

1) None of the segments of the different models are assigned to the same segment of the object.

2) One or more segments of the models which do not match to any of the segments of the object should be assigned to the nil class. This condition is taken care of in our formulation of the problem in section II.

The total criterion of consistency and ambiguity for all the M models is,

$$F(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_M) = \sum_{m=1}^{M} \sum_{i=1}^{N_m} J_{im}^{(n)}(\vec{v}_m), \quad n = 1,2 \tag{6}$$

where N_m is the number of segments, $\vec{v}_m$ is the $N_m L$ dimensional probability vector associated with the mth model X_m, and $J_{im}^{(n)}(\vec{v}_m) = \vec{p}_{im} \cdot \vec{q}_{im}^{(n)}$. The occlusion condition 1) is,

$$G(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_M) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} g(\vec{s}_i,\vec{s}_j) = 0 \tag{7}$$

where $\vec{s}_\ell$ is obtained from $\vec{v}_\ell$ with the elements corresponding to the nil class set equal to zero for all the units of the model X_ℓ and $g(\vec{s}_i,\vec{s}_j)$ is the inner product of the vectors $\vec{s}_i$ and $\vec{s}_j$.

The penalized objective function [3] is,

$$\psi_c(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_M) = F(\vec{v}_1,\vec{v}_2,\dots,\vec{v}_M) + \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} d_{ij}\phi_{ij}[g(\vec{s}_i,\vec{s}_j)] \tag{8}$$

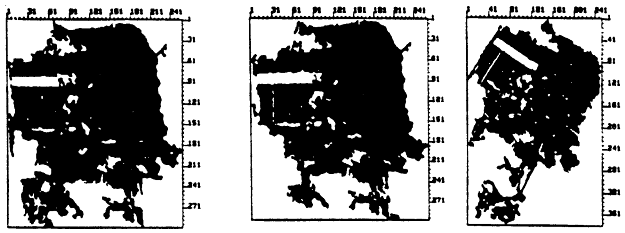
where ϕ_{ij} is a penalty function and $\{d_{ij}\}$ are penalty constants. The penalty function is taken as the simple quadratic loss function. Now the problem becomes equivalent to that of maximizing (8) subject to the linear probability constraints.

IV Examples

Example 1 : Fig. 1 shows three regions corresponding to the same area in two San Francisco images which are rotated with respect to each other. The approximation of the regions reduced by a factor of 14 is shown in Fig. 2. We want to match the shape of objects (Fig. 2(a) and (b)) against the model (Fig. 2(c)) so object segments are units in this example. Note that segments 13 and 17 of object 1 match with segments 7 and 14 or 16 of the model respectively. Similarly segments 10 and 15 or 16 of object 2 match with the segments 7 and 14 or 16 of the model. Results of shape matching are shown in Table 1. Most of the assignments are very reasonable and correct. Label 30 is the nil class. Using the results of labeling the relative rotation between the object 1 and the model is found to be 36.1° and between the object 2 and the model it is 35.5°. The actual rotation is about 35°.

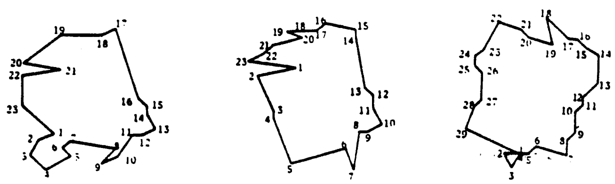
Example 2 : Figure 3 shows 512x512, 8 bit images of occluding industrial parts. The images in figs. 3(a) and (b) are reduced by 16 times and the image in fig. 3(c) by 18 times. The polygonal approximation is shown in fig. 4. Only the rotation and scale invariant features are used. Label 25 is the nil class. The results of the occlusion algorithm are shown in Table 2. Note that all the key assignments of the units are correct.

The technique is found to be very effective when applied to the aerial, industrial and bilological images [2]. The computation time varied from 4 seconds to 5 minutes on a PDP-10.



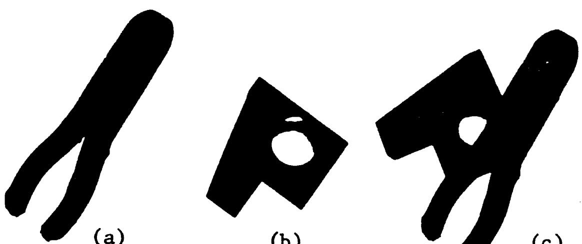
(a) (b) (c)

Figure 1. Golden Gate Park Region obtained from two San Francisco Images using a recursive region splitting technique. Regions shown are at different scales.



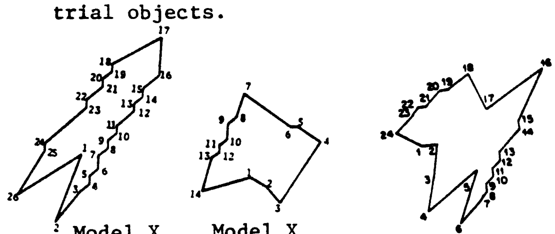
object 1 object 2 Model

Figure 2. Polygonal approximations of the regions shown in Figure 1.



(a) (b) (c)

Figure 3. Partial occlusion of two industrial objects.



Model X_1 Model X_2 Apparent object

Figure 4. Polygonal approximation of the industrial objects of Figure 3.

REFERENCES

[1] O.D. Faugeras and M. Berthod, "Improving Consistency and Reducing Ambiguity in Stochastic Labeling: An Optimization Approach," IEEE Trans. on PAMI, Vol. PAMI-3, pp. 412-424, July 1981.

[2] B. Bhanu, "Shape Matching and Image Segmentation Using Stochastic Labeling," Ph.D Thesis, USCIPI Report 1030, Image Processing Institute, USC, Los Angeles, August 1981.

[3] D.G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley Publishing Co., 1973.

Table 1. Label of Segments, Example 1

Segments of the Object 1,2	Labels at Different Iterations							
	Object 1				Object 2			
	0	5	2	7	0	6	1	6
1	30	27	27	27	1	1	1	1
2	28	28	28	28	2	2	2	2
3	25	29	28	28	30	27	27	3
4	5	3	3	3	28	28	28	28
5	30	5	5	5	30	29	29	29
6	30	4	5	5	30	2	2	2
7	6	6	6	6	3	3	3	3
8	19	1	7	7	4	4	4	4
9	30	3	5	5	30	5	5	5
10	5	3	3	3	7	7	7	7
11	4	4	4	4	30	8	8	8
12	30	5	5	5	11	9	9	9
13	7	7	7	7	30	12	12	11
14	30	8	8	8	30	13	14	14
15	30	9	9	9	14	14	14	14
16	17	17	13	13	30	16	16	16
17	30	30	18	16	30	17	19	19
18	30	20	20	20	30	19	19	19
19	30	30	23	23	18	18	18	18
20	18	18	30	30	1	19	19	19
21	1	21	19	19	30	22	20	20
22	22	22	22	22	23	23	23	22
23	30	26	26	26	18	29	29	29
Value of Criteria	—	2.2	1.6	1.7	--	1.9	1.6	1.7
	$J^{(1)}$		$J^{(2)}$		$J^{(1)}$		$J^{(2)}$	

Table 2. Results of Labeling for the Model X_1 and X_2, Example 2

Units of Model X_1&X_2	Labels at Different Iterations							
	Model X_1				Model X_2			
	First Stage		Second Stage		First Stage		Second Stage	
	0	3	1	5	0	3	1	5
1	5	5	5	5	25	25	25	25
2	6	6	6	6	25	25	25	23
3	25	7	25	7	25	25	25	25
4	25	8	8	8	25	25	25	25
5	25	25	25	25	25	25	25	25
6	25	25	25	25	25	25	25	15
7	25	25	25	25	25	18	18	18
8	25	25	25	25	25	25	25	25
9	25	25	25	25	25	25	25	25
10	25	25	25	25	25	25	25	25
11	25	25	25	25	25	25	25	25
12	25	25	25	25	25	25	25	19
13	25	25	25	25	25	25	25	22
14	25	25	25	25	25	25	25	25
15	25	25	25	25				
16	25	25	25	14				
17	25	25	25	25				
18	25	25	25	25				
19	25	25	25	25				
20	25	25	25	25				
21	25	25	25	25				
22	25	25	25	25				
23	25	25	25	25				
24	14	14	14	14				
25	25	25	25	25				
26	25	4	25	25				
First Term	--	4.7	16.2	17.2	--	2.8	5.6	7.1
Penalty Term	--	.47	1.6	0	--	.28	.56	0
Criterion	--	4.2	14.6	17.2	--	2.5	5.1	7.1
Penalty Const.	--	.008	1.7	--	--	.005	.59	--

Bir Bhanu
June 13, 1982

proceedings

PR IP 82

IEEE Computer Society
Conference on

pattern recognition and image processing

82CH1761-6
82-80366

C PUTER
S ETY
PR SS

IEEE COMPUTER SOCIETY

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

Bir Bhanu

proceedings

PR IP 82

JUNE 14–17, 1982
LAS VEGAS, NEVADA

IEEE Computer Society
Conference on

pattern recognition and image processing

82CH1761-6
82-80366

Additional copies may be ordered from
IEEE Order Dept.
445 Hoes Lane
Piscataway, NJ 08854

IEEE COMPUTER SOCIETY PRESS

IEEE Computer Society
10662 Los Vaqueros Circle
Los Alamitos, CA 90720

IEEE COMPUTER SOCIETY

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.