BINARY SHAPE RECOGNITION BASED ON AN AUTOMATIC MORPHOLOGICAL SHAPE DECOMPOSITION

Yunxin Zhao
Robert M. Haralick

Department of Electrical Engineering, FT-10
University of Washington
Seattle, WA 98195
USA

Abstract — This paper presents a technique for translation invariant binary convex polygon shape recognition based on a morphological shape decomposition. The triangle shape primitives from the decomposition of convex shapes are used as features for shape recognition. The shape primitives are smaller and simpler than the templates of shapes, thus they are more efficient in representing shapes for discrimination. Maximum entropy reduction is used as an optimization criterion for selecting features from among the shape primitives at each node of a decision tree. Experiments on the classification of 10 classes of noisy polygon shapes, where 5 replications per class were used for training and 50 replications per class were used for testing, achieved a recognition rate of 98.80% on the test set.

I. INTRODUCTION

Mathematical morphology is a natural processing approach for image object identification since it is a technique based directly on shape [1]. Shape recognition using morphology demands systematical methods of selecting features from a given set of shapes to be classified. The problem approached in this paper is based on the morphological decomposition of convex shapes [2,3]. As an inherited characteristic from morphological operations, the recognition technique is shape-translation invariant. The recognition system is a decision tree classifier. The shape primitives from the decomposition are used as features at nonterminal nodes. The appeal of this approach lies in the fact that the shape primitives are smaller and simpler than the templates of shapes, and they can be generated automatically using the decomposition technique. For efficient classification, maximum entropy reduction is used as a criterion in selecting features from among all shape primitives [4,5]. Although the system works on convex polygon shapes, it is of practical value since complex shapes can be decomposed into their convex pieces [6]. Thus the algorithm described here can be those for a subsystem of a complex shape recognition system.

This paper is organized into 6 sections. In Section 2, the background materials on morphological operations and decision tree classifiers are provided. The basic ideas of recognizing shapes using morphological shape decomposition is given in Section 3. The algorithm for the construction of the decision tree is in Section 4. In Section 5 the experimental results are provided, and Section 6 contains a concluding discussion.

II. BACKGROUND

Morphological Operations

Morphological operations are defined on sets in E^N. Sets in E^2 correspond to the foreground regions in binary images. Let $P \in E^2$ be the image undergoing analysis, $S \in E^2$ be the structuring element, and let the translation of P by $s \in S$ be denoted by $(P)_s$, *i.e.*,

$$(P)_s = \{ \hat{p} \in E^2 \mid \hat{p} = p + s \text{ for } some\ p \in P \}.$$

The morphological operation of dilation, erosion, and opening of P by S are defined as follows [1]:

Dilation: $P \oplus S = \bigcup_{s \in S} (P)_s$,

Erosion: $P \ominus S = \bigcap_{s \in S} (P)_{-s}$,

Opening: $P \circ S = (P \ominus S) \oplus S$.

From shape transformation point of view, dilation is a shape-expansion operation, erosion is a shape contraction operation, whereas opening is a compound shape transformation defined as the erosion of an image followed by the dilation of the eroded result. The opening of an image corresponds to an elimination of foreground image details smaller than the structuring element without introducing global geometric distortion of unsuppressed features. If P is unchanged by opening it with S, we say P is totally openable by S; if the opening results in an empty set, then we say P is not openable by S.

Decision Tree Classifier

The decision tree classifiers are well suited to a feature based recognition system, where successively more features can be utilized to discriminate more and more detailed differences between objects through a hierarchical decision procedure. Theoretical guidance to the design of a decision tree classifier is available in [4], where a purity function describing classification purity was defined as a quantitative measure for optimizing the threshold of decision rules at each nonterminal node. The maximization of the purity function is shown in [7] as including the maximization of the entropy reduction in a recognition system as a special case. In this paper we use the entropy reduction as the criterion for selecting features at each nonterminal node.

The 1st author is currently with Speech Technology Lab., Division of Panasonic Technologies, Inc., 3888 State Street, Santa Barbara, CA 93105.

We consider a non-overlapping decision tree classifier. Assuming that at the root node there are N classes of objects and each class has the same number of training samples. A decision rule is applyed at the root node to split the training samples into descendent child nodes, each corresponding to a decision region. For a non-overlapping decision tree, the decision boundaries are constrained such that no training samples of the same class will be split into different child nodes. If at any child node, there are training samples from more than one class, a decision rule at that node is then applied to further split the training samples into the children nodes of that node; if a child node has training samples from only one class, then it becomes a terminal node. This process is repeated until N terminal nodes are generated, each being associated with one class. Specifically, consider the jth node of the decision tree, let $N^{(j)}$ be the number of possible classes for a training sample at the node, h_j be the number of decision regions produced by the decision rule at the node, $n_q^{(j)}$ be the number of shape classes in the qth decision region, the entropy reduction due to the expansion of the node j into the h_j child nodes is then defined as [7,5]

$$\Delta H^{(j)} = \log N^{(j)} - \sum_{q=1}^{h_j} \frac{n_q^{(j)}}{N^{(j)}} \log n_q^{(j)} .$$

This measure of entropy reduction is used in the work of this paper.

III. SHAPE DISCRIMINATION

Decomposition of Convex Polygon Shapes

It was shown in [2] that any convex polygon shape in the real plane R^2 is decomposable through dilation into triangles (line segments are considered as special cases of triangles). Let P be a convex polygon shape in R^2, and S_i, $1 \le i \le n$, be the triangles from the decomposition of P, then P can be represented as the successive dilations of these S_i's, *i.e.*,

$$P = S_1 \oplus S_2 \cdots \oplus S_n .$$

We call the triangles S_i, $1 \le i \le n$, the shape primitives of P. Fig. 1 illustrates the decomposition of a polygon P into a line segment S_1 and three triangles S_2, S_3, and S_4. For details about the decomposition procedure we refer readers to [2].

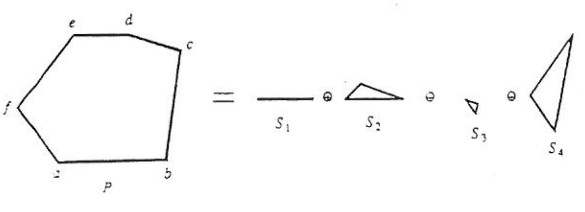

Fig. 1 A convex polygon P is decomposed into 4 shape primitives S_1 through S_4.

To decompose a polygon shape P in the discrete plane Z^2, we first map P to the real plane R^2: $P \rightarrow C(P)$, then decompose $C(P)$ into $C(S_1)$, $C(S_2)$, $\cdots$, $C(S_n)$, and finally we map the shape primitives back to the Z^2 plane: $C(S_i) \rightarrow S_i$, $i = 1, \cdots, n$. The mapping of a polygon shape from Z^2 to R^2 is taken as a convex hull transformation; and the mapping of the shape primitives from R^2 to Z^2 is a sampling on the image grid points. For some problems of the sampling effect on shape decomposition see [2,7] for details.

Discriminating Shapes Using Shape Primitives

The idea of using shape primitives for shape discrimination is based on the following pair of propositions [1]:

(1). $P \circ S \subseteq P$

(2). *If* $P = S_1 \oplus S_2$, *then* $P \circ S_2\,(or S_1) = P$.

This pair of propositions indicate that if a shape is decomposable through dilation into a set of shape primitives, then it is totally openable by any of its shape primitives, but it is not openable or not totally openable by shape primitives that do not belong to any of its decomposition. When we refer that a shape primitive belongs to a decomposition, we also imply those that are similar to the shape primitive but are of smaller sizes. The normalized opening residue measures the degree of matching of a shape P to a shape primitive S, *i.e.*,

$$r = 1 - \#(P \circ S) / \#(P) ,$$

where $\#(X)$ means the number of pixels in the shape X having value of 1. If S is a shape primitive of P, then $r = 0$; otherwise $0 < r \le 1$. The normalized opening residue can thus be used as a discriminant function in the decision tree construction.

The distance between opening residues of distinct shape classes can further be improved by using the next proposition [7]:

(3) $P \circ (S_1 \oplus S_2) \subseteq P \circ S_1$.

This proposition together with proposition (2) indicate that if $S_{1,1}$ and $S_{1,2}$ are shape primitives for shape P_1 but are not shape primitives for shape P_2, then the difference of opening residues between P_1 and P_2 will increase by using the structuring element $S_1 \oplus S_2$ than using S_1 or S_2 alone. Therefore if the distance between opening residues of different shape classes are not large enough, the feature structuring element can be enlarged by the dilation of another shape primitive.

Selecting Features from Shape Primitives

Given the decomposed shape primitives of the set of shapes to be classified, a certain criterion needs to be developed to select the important and discriminating shape primitives as features for shape recognition. It is commonly true that large sized shape primitives represent the main feature of a shape whereas smaller ones represent details. In the classification stage, using a large sized shape primitive to open the test shapes can immediately eliminate the possibility of any smaller sized shapes and hence reduce the search complexity. The importance of the shape primitives can therefore be determined according to their sizes. For triangles the perimeter lengths are used here for size description. As this notion of size is mainly empirical, the maximum entropy reduction is further used as a quantitative measure of the importance of the shape primitives. Considering the tradeoff between optimality and training time, the entropy reduction is calculated for a few large-sized shape primitives at a node and the one giving maximum entropy reduction is selected as the feature structuring element. Taking into account of the effect of noise, the size of shape primitives are also varied from its decomposed value to some smaller scale. This size variation is also treated as a parameter to be optimized using the entropy criterion.

IV. DECISION TREE CONSTRUCTION

Assuming that the training shape samples are from N classes, each class has γ samples, and for each class a representative shape can be obtained from the γ samples, the construction of a decision tree involves the following steps.

Step 1.

Decompose the representative shapes for each class. Order the shape primitives of each shape class from large to small according to their perimeter lengths. Specifically, let $S_k(m;\lambda)$ denote the mth shape primitive (its sides are scaled by the scale parameter λ) of the representative shape of class k, the shape primitive sets for the N representative shapes are:

$$\{S_i(1;\lambda),\ S_i(2;\lambda),\ \cdots\ S_i(m_1;\lambda)\}\ ,\ i = 1, \cdots, N.$$

where $\lambda \in \Lambda$, and Λ contains a few candidate scaling factors with $\max(\lambda) = 1.0$.

Step 2.

At each node undergoing expansion, we have a set of training shape samples and a set of structuring elements which are shape primitives chosen as candidates of the feature structuring element for the node. Let $P_{i,v}^{(j)}$ be the vth training sample from shape class i at node j, assuming the samples are from $N^{(j)}$ classes, then the set of shape samples is $\{\ P_{i,v}^{(j)},\ 1 \leq v \leq \gamma\ ,\ 1 \leq i \leq N^{(j)}\}$. Let $S_k^{(j)}(1;\lambda)$ denote the 1st shape primitive, scaled by λ, of the kth class at the node, the structuring element set is $\{S_k^{(j)}(1;\lambda),\ \lambda \in \Lambda,\ k \in K\ \}$, where K contains the indices of the largest few shape primitives at the node, $|K|$ is kept small considering computation time.

For each shape sample $P_{i,v}^{(j)}$, calculate the residue $r_{i,v,k}^{(j)}(1;\lambda)$'s by opening the shape with $S_k^{(j)}(1;\lambda)$ for every $\lambda \in \Lambda$ and $k \in K$. The residue is defined as

$$r_{i,v,k}^{(j)}(1;\lambda) = 1 - \#(P_{i,v}^{(j)} \circ S_k^{(j)}(1;\lambda)) / \#(P_{i,v}^{(j)})\ .$$

Step 3.

For each shape primitive $S_k^{(j)}(1;\lambda)$, formulate an opening residue interval $R_{i,k}^{(j)}(\lambda)$ for each shape class i, $1 \leq i \leq N^{(j)}$, where $R_{i,k}^{(j)}(\lambda)$ is defined as the coverage of the residue intervals of the class, *i.e.*,

$$R_{i,k}^{(j)}(\lambda) = [\ \min_{1 \leq v \leq \gamma}\{\ r_{i,v,k}^{(j)}(\lambda)\ \},\ \max_{1 \leq v \leq \gamma}\{\ r_{i,v,k}^{(j)}(\lambda)\ \}].$$

Step 4.

Expand the node j into child nodes. The child nodes are formed by assigning the shape classes with their residue intervals separated smaller than a margin ϵ_r into the same decision region, see [7] for details.

Assuming h_j number of child nodes are thus generated, there are then h_j residue intervals, one for each node, which is the coverage of the residue intervals of every shape class contained in the node. Let $\Omega_{i,k}^{(j)}(\lambda)$ denote the residue interval for the ith child node, $1 \leq i \leq h_j$, where the child nodes are indexed by the order that their residue intervals occur on the real axis from left to right.

Step 5.

Let $n_{q,k}^{(j)}(\lambda)$ be the number of shape classes in the qth child node, the entropy reduction due to the expansion of node j using the structuring element $S_k^{(j)}(1;\lambda)$ is then

$$\Delta H_k^{(j)}(\lambda) = \log N^{(j)} - \sum_{q=1}^{h_j} \frac{n_{q,k}^{(j)}(\lambda)}{N^{(j)}} \log n_{q,k}^{(j)}(\lambda)\ .$$

The $\Delta H_k^{(j)}(\lambda)$ is calculated for all possible $k \in K$ and $\lambda \in \Lambda$. Let δ be the minimum entropy reduction of interest, If $\Delta H_{k^*}^{(j)}(\lambda^*) = \max_{k,\lambda} \Delta H_k^{(j)}(\lambda) \geq \delta$, then $S_{k^*}^{(j)}(\lambda^*)$ is chosen as the feature structuring element for node j. The threshold between the decision regions of the mth and the $m+1$th child nodes is calculated as

$$TH(m) = (\max\{\Omega_{m,k^*}^{(j)}(\lambda^*)\} + \min\{\Omega_{m+1,k^*}^{(j)}(\lambda^*)\})/\ 2\ .$$

If $\max_{k,\lambda} \Delta H_k^{(j)}(\lambda) < \delta$, then change the structuring element list at the node to

$$\{\ S_k^{(j)}(1;\lambda) \oplus S_k^{(j)}(2;\lambda),\ \lambda \in \Lambda,\ k \in K\ \},$$

i.e., each structuring element is enlarged through the dilation of another shape primitive. Using the list of new structuring elements as candidates for the feature structuring element, repeat the steps 2 through 5 on the same node until a satisfactory entropy reduction is obtained, see [7] for details.

Step 6.

Repeat the steps 2 through 5 on every node waiting to be expanded until N terminal nodes are produced, each being associated with a shape class.

V. EXPERIMENT

Experiments were performed on random polygons of 10 classes, the average block size of the polygons was 131×121. The vertices of the polygons were disturbed by a uniformly distributed noise on the interval [-5, 5]. A disk shape noise were further added to the boundaries of the vertex-disturbed polygons. The centers of the disk noise were at each pixel on the edges of a polygon and their radius were varied over [0, 3]. This noise transforms the smooth edges of a polygon into zigzag ones.

In training, 5 noisy shapes from each class were used. A representative shape for a class is obtained by averaging the corresponding vertices from the 5 vertex-disturbed shapes. This representative one is then decomposed into a set of triangles. The 10 representative shapes are shown in Fig. 2. In constructing the decision tree, the residue margin ϵ_r for shape class separation was chosen as 0.2, the minimum entropy reduction δ was 0.0, the number of shape primitives considered at each node, $|K|$, was 2, and the scaling factor $\lambda \in \Lambda = \{1.0, 0.9, 0.8\}$. The decision tree constructed is shown in Fig. 3. Each node of the tree displays a list of shapes to be classified, the feature structuring element used in expanding the node, and the entropy reduction (in natural logarithmic) thus obtained. Each branch of the tree displays the threshold of residue that classifies the shapes into the nodes under the branch. The feature structuring elements in the decision tree are shown in Fig. 4. In recognition, 50 noisy shapes (independently generated with the same noise level used in the training sample) for each

10 classes were tested and the results are tabulated in Table 1. Of the 500 noisy shapes 494 were classified correctly. The system achieved a recognition rate of 98.80%.

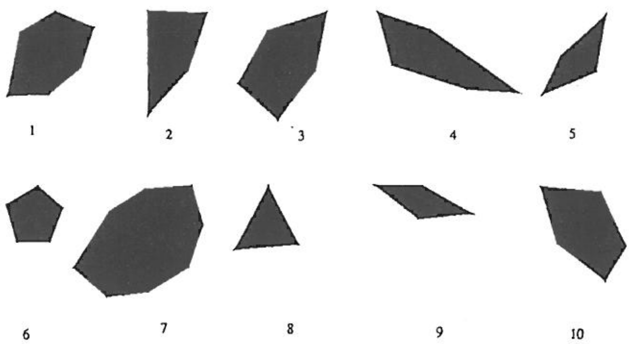

Fig. 2. The representative shapes from the 10 shape classes used in the experiment.

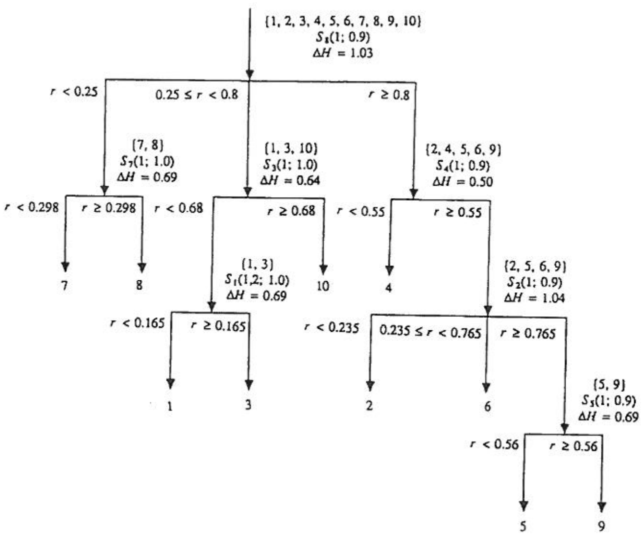

Fig. 3. The classification tree: each node displays a list of shapes to be classified, the structuring element used in expanding the node, and the entropy reduction (in natural logarithmic) thus obtained; each branch displays the residue threshold that classifies the shapes into the nodes under the branch.

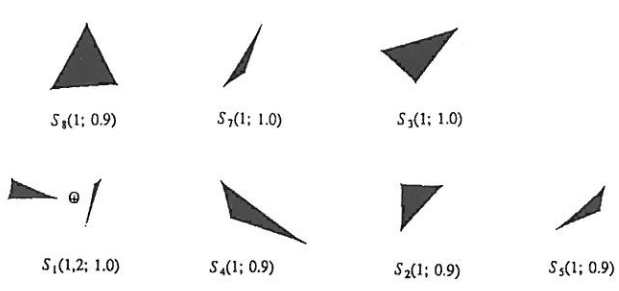

Fig. 4. The feature shape primitives used in the classification tree.

Table 1. Recognition result

test shape	1	2	3	4	5	6	7	8	9	10
1	47	0	3	0	0	0	0	0	0	0
2	0	50	0	0	0	0	0	0	0	0
3	0	0	50	0	0	0	0	0	0	0
4	0	0	0	50	0	0	0	0	0	0
5	0	0	0	0	48	2	0	0	0	0
6	0	0	0	0	0	50	0	0	0	0
7	0	0	0	0	0	0	50	0	0	0
8	0	0	0	0	0	0	0	50	0	0
9	0	0	0	0	0	0	0	0	50	0
10	0	0	1	0	0	0	0	0	0	49

VI. CONCLUDING DISCUSSION

The experiment showed that simple shape primitives from the morphological decomposition of shapes have potential power for shape discrimination. The shape primitives are much smaller than the templates of shapes, thus a considerable saving in decision time is achieved. We believe the recognition errors occurring in this experiment were mainly due to the insufficient number of shape samples used in training. The calculated residue threshold may not always have been optimal. Greater flexibility in the decision tree could be obtained if we used the overlapping rather than the non-overlapping decision tree [4]. For systems working on a large number of shape classes, some savings in work can occur by measuring the similarity between shape primitives. Searching for features among the large set of shape primitives can then be reduced to a searching among similar shaped shape primitive classes.

ACKNOWLEDGEMENT

The authors would like to thank some discussions of Prof. X. Zhuang during his visiting the department.

REFERENCES

[1] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image Analysis Using Mathematical Morphology", *IEEE Trans. PAMI,* **PAMI-9**, pp. 532-550, May, 1986.

[2] P. K. Ghosh and R. M. Haralick, "Morphological Decomposition of Binary Image," in preparation.

[3] X. Zhuang and R. M. Haralick, "Morphological Structuring Element Decomposition," *Computer Vision, Graphics and Image Processing,* **35**, pp. 370-382, 1986.

[4] R. M. Haralick and L.G. Shapiro, **Computer and Robot Vision** , to be published by Addison-Wisley, 1989.

[5] Q. Wang and C. Suen, "Analysis and Design of a Decision Tree Based on Entropy Reduction and Its Application to Large Character Set Recognition," *IEEE Trans. PAMI,* **PAMI-6**, pp. 406-417, July, 1984.

[6] T. Pavlidis, **Structural Pattern Recognition.** Springer Publ., New York, 1977.

[7] Y. Zhao, P. K. Ghosh and R. M. Haralick, "Binary Shape Recognition Based on an Automatic Morphological Shape Decomposition," in preparation.