Two-Dimensional Object Recognition Through Two-Stage String Matching

Wen-Yen Wu and Mao-Jiun J. Wang

***Abstract*—A two-stage string matching method for the recognition of two-dimensional (2-D) objects is proposed in this work. The first stage is a global cyclic string matching. The second stage is a local matching with local dissimilarity measure computing. The dissimilarity measure function of the input shape and the reference shape is obtained by combining the global matching cost and the local dissimilarity measure. The proposed method has the advantage that there is no need to set any parameter in the recognition process. Experimental results indicate that the two-stage string matching approach significantly improves the recognition rates while comparing to the one-stage string matching method.**

***Index Terms*— Compactness, cost function, cyclic string, edit graph, feature extraction, object recognition, two-stage string matching.**

I. Introduction

Recognizing two-dimensional (2-D) objects is very important in many applications. It has been agreed that representation and matching are the two major problems involved in pattern recognition. To solve these problems, many methods have been proposed [4]. Some examples of the matching techniques are template matching, string matching, shape-specific point matching, principal axis matching, dynamic programming, mutually-best matching, chamfer matching, graph matching, relaxation, and elastic matching, etc. The 2-D object recognition techniques can be classified into two major categories: the statistical method and the syntactic method. Due to the advantages and disadvantages of these two methods are complementary, a combined approach has been proposed [1], [3], [5], [6], [9], [12].

Suppose that $\boldsymbol{s}$ and $\boldsymbol{t}$ are two strings. Further, given an *edit cost function*, three types of arcs which represent the insertion, deletion, and change operations, respectively, can be defined. Wagner and Fischer [9] defined the *edit graph associated with* $\boldsymbol{s}$ *and* $\boldsymbol{t}$ by the weighted graph $\boldsymbol{G}$ (see Fig. 1) with vertices $v(i, j) = 0, 1, \cdots, n, j = 0, 1, \cdots, m$. The problem of finding a *minimum cost edit sequence* taking $\boldsymbol{s}$ to $\boldsymbol{t}$ is now equivalent to find a *shortest path* in $\boldsymbol{G}$ from $v(0, 0)$ to $v(n, m)$. The above method solves the *linear string-to-string correction problem* by finding the edit distance and its corresponding edit sequence.

Tsai and Fu [6] have tried to introduce the statistical decision theory into the attributed grammar such that it might be more practical for applications. They concluded that the attributed grammar is a good tool for combining statistical and syntactic methods. Tsai and Yu [7] proposed a powerful edit operation, which is the merge operation. With the merge operation, the recognition rates can be increased. Further, Tsay and Tsai [8] introduced another new edit operation, i.e., split, to the attributed string matching.

It needs to define a reference line for the linear string matching techniques. To overcome this problem, Maes [3] proposed a cyclic string matching technique for polygonal shape recognition. He used

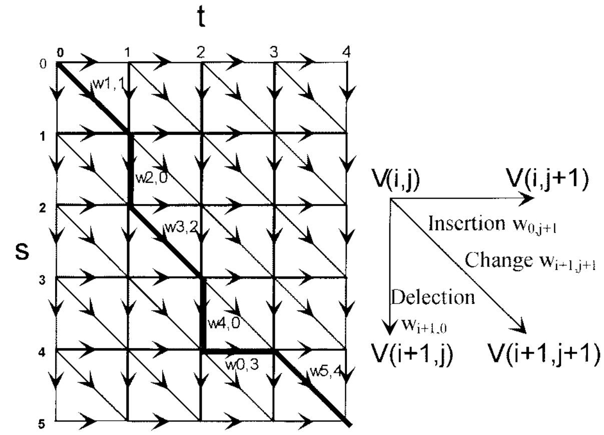


Fig. 1. Edit graph $\boldsymbol{G}$ for $|\boldsymbol{s}| = 5$ and $|\boldsymbol{t}| = 4$, and the shortest path.

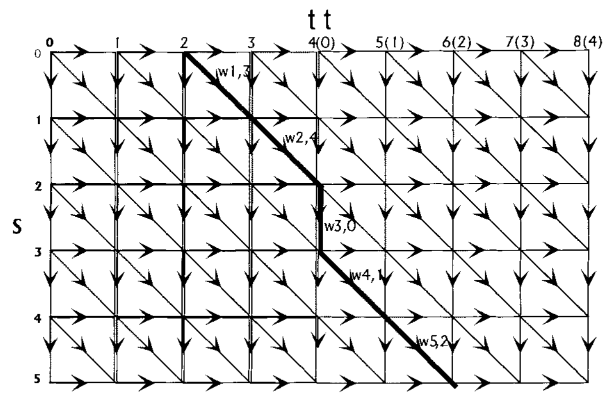


Fig. 2. Edit graph $\boldsymbol{H}$ associated with $\boldsymbol{s}$ and $\boldsymbol{tt}$, and the shortest path.

cyclic strings to represent polygons. Maes first constructed the edit graph $\boldsymbol{H}$ associated with $\boldsymbol{s}$ and $\boldsymbol{tt}$ (see Fig. 2), where $\boldsymbol{tt} = t_1 t_2 \cdots t_m t_1 t_2 \cdots t_m$ is the string which concatenates $\boldsymbol{t}$ with itself. It becomes the *cyclic string-to-string correction problem*. The minimum edit distance can be found and its corresponding edit sequence can be identified [2]. Only the conventional three edit operations were needed for the cyclic string matching.

Besides, Maes [3] suggested two main directions to the string matching problem affected by uneven segmentations. The first approach is to develop a polygonal approximation algorithm, which is free from uneven segmentation. Since the edit sequence often provides a correct matching of the significant features between the two approximated polygons, the second approach is to use the result of matching and conduct a more detailed analysis on these two polygons. Based on the latter suggestion, we propose a new two-stage string matching method.

II. Object Recognition by Two-Stage String Matching

In the first stage of matching, a global cyclic string matching is conducted to find the best-matched pair between the input shape and the reference shapes. The second stage of matching is followed to superimpose the input shape on each reference shape and to compute the local dissimilarity measure between the input shape and each reference shape. Finally, the dissimilarity measure between the input

Manuscript received February 7, 1997; revised September 10, 1998. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dmitry B. Goldgof.

W.-Y. Wu is with the Department of Industrial Management, I-Shou University, Kaohsiung, Taiwan 84008, R.O.C. (e-mail: wywu@csa500.isu.edu.tw).

M.-J. J. Wang is with the Department of Industrial Engineering, National Tsing Hua University, Hsinchu, Taiwan 30043, R.O.C.

Publisher Item Identifier S 1057-7149(99)05106-4.

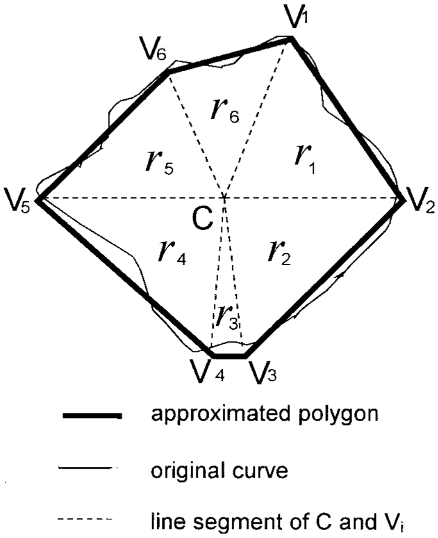


Fig. 3. Illustration of the computing of the reciprocal of compactnesses $r_i = a_i/p_i^2$, for $i = 1, 2, \cdots, 6$, where V_i's are the dominant points and C is the centroid of the object.

shape and each of the reference shapes is then obtained by combining the global matching cost and the local dissimilarity measure in the first and second stages, respectively. The input shape is matched to the reference shape that has the minimum dissimilarity measure among all the reference shapes.

A. Global Cyclic String Matching

The normalized reciprocal of compactnesses is used as the global feature in the string matching process [11]. Suppose that V_i's, for $i = 1, 2, \cdots, N$, are the dominant points of object [10] and C is the centroid of the object (see Fig. 3). The reciprocal of the compactnesses of the triangles formed by two adjacent dominant points and the centroid of the object can be obtained. Let r_i be the ith reciprocal of the compactness, which is defined as

$$r_i = \frac{a_i}{p_i^2}, \quad \text{for } i = 1, 2, \cdots, N \tag{1}$$

where $p_i = |\overline{V_iV_{i+1}}| + |\overline{V_iC}| + |\overline{V_{i+1}C}|$ is the perimeter and a_i is the area of the triangle.

Let $\boldsymbol{I}$ and $\boldsymbol{R_i}$ represent the approximated polygons of input shape and the ith reference shape, for $i = 1, 2, \cdots, c$, where c is the number of reference shapes, respectively. Further, let n and m be the numbers of vertices for $\boldsymbol{I}$ and $\boldsymbol{R_i}$, respectively. Now, the reciprocal of the compactnesses are denoted as the *symbols* and we can use the *strings* $\boldsymbol{s} = s_1 s_2 \cdots s_n$ and $\boldsymbol{t} = t_1 t_2 \cdots t_m$ to represent the input shape and the reference shape, respectively, [3]. For convenience, the string with zero length is called the *null string*, and it is denoted as λ.

Let $\boldsymbol{X}$ be the set which the elements are the reciprocal of the compactnesses, i.e., s_i or t_j. Since the reciprocal of compactnesses and the vertices of polygons can be represented as cyclical, the problem of matching two shapes is therefore identical to the cyclic string matching between strings $\boldsymbol{s}$ and $\boldsymbol{t}$. Given an edit cost function ϵ, then we can construct the edit graph $\boldsymbol{H}$ associated with $\boldsymbol{s}$ and $\boldsymbol{tt}(= t_1 t_2 \cdots t_m t_1 t_2 \cdots t_m)$ to find the shortest path from the left-upper corner to the right-down corner (see Fig. 2). By tracing the minimum cost edit sequence, the matching relation between the vertices of $\boldsymbol{I}$ and $\boldsymbol{R_i}$ can be determined (Fig. 4). For instance, the two ordered sequences $\boldsymbol{B_{I,i}} = (\boldsymbol{I}_{1i}, \boldsymbol{I}_{2i}, \cdots, \boldsymbol{I}_{ki})$ and $\boldsymbol{B_{R,i}} = (\boldsymbol{R}_{1i}, \boldsymbol{R}_{2i}, \cdots, \boldsymbol{R}_{ki})$ represent that the $\boldsymbol{I}_{ji}$th vertex of $\boldsymbol{I}$ matches with the $\boldsymbol{R}_{ji}$th vertex of $\boldsymbol{R_i}$, for $i = 1, 2, \cdots, k$, where k is the number

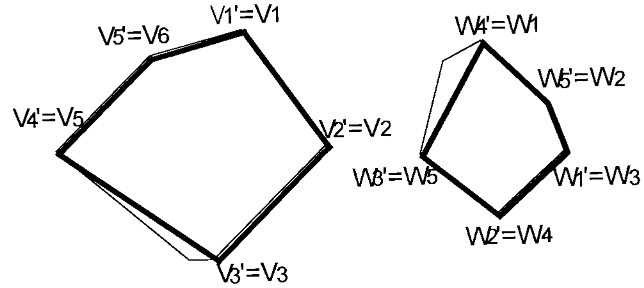


Fig. 4. Illustration of determining the matched vertices from the best-matched pair, where $\boldsymbol{B}_{I,i}$ = (1, 2, 3, 5, 6) and $\boldsymbol{B}_{R,i}$ = (3, 4, 5, 1, 2).

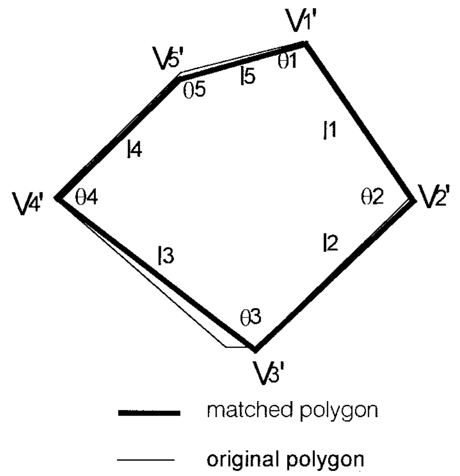


Fig. 5. Length $l_i = |\overline{V_i'V_{i+1}'}|$ and the angle $\theta_i = \angle V_{i-1}'V_i'V_{i+1}'$, for $i = 1, 2, \cdots, 5$, where V_i' is the ith matched dominant point.

of matched points. The two ordered sequences are called the *best-matched pair* of $\boldsymbol{I}$ and $\boldsymbol{R_i}$. And the minimum edit distance in the global string matching stage is called *matching cost* [3].

Further, suppose that the value of λ is zero. Then, for all $s_i, t_j \in \boldsymbol{X}$, the *edit cost function* ϵ can be simply defined as

$$\epsilon(s_i \to t_j) = |s_i - t_j|. \tag{2}$$

We do not have the problem of choosing proper weighting factors since only one feature is used in our global string matching.

B. Local Matching for Dissimilarity Measures Computing

Based on the information of the best-matched pair in the first stage, we can assess the local dissimilarity measure between the input shape and the reference shape by their corresponding lengths and angles. Let V_i', for $i = 1, 2, \cdots, k$, be the ith matched dominant point, where k is the number of matched points. Suppose that l_i and Θ_i are the length of $\overline{V_i'V_{i+1}'}$ and the angle of $\angle V_{i-1}'V_i'V_{i+1}'$, respectively, (Fig. 5).

The local dissimilarity measure computation is a straightforward task. The best-matched pair obtained in the global string matching stage gives a one-to-one mapping between the vertices of the two polygons. The lengths and angles are then computed. The local dissimilarity measure between the two polygons is determined by averaging the absolute differences of the lengths and the angles. Therefore, the local dissimilarity measure between the input shape and the reference shape can therefore be obtained by

$$\mathrm{LD} = \frac{1}{2k}\left(\sum_{i=1}^{k} |l_i - \overline{l_i}| + \sum_{i=1}^{k} |\theta_i - \overline{\theta}_i|\right) \tag{3}$$

where k is the number of matched vertices, and $l_i(\overline{l_i})$ and $\theta_i(\overline{\theta}_i)$ are the ith normalized length and normalized angle of the input (reference) shape, respectively.

It is better to use both the global and local information in assessing the dissimilarity measure between two objects. The matching cost obtained in the global string matching stage represents the global dissimilarity measure of the input shape and the reference shape. Suppose that the matching cost in the global string matching stage is GD. The dissimilarity measure between the input shape and the reference shape can be defined as a function of LD and GD.

$$D = f(\mathrm{LD}, \mathrm{GD}). \tag{4}$$

Suppose that there are c classes of reference shapes. The dissimilarities between the input shape and each of the reference shapes can be obtained. Let D_i be the dissimilarity measure function between the input shape and the ith reference shape, for $i = 1, 2, \cdots, c$. The input shape can then be classified as the reference shape with the minimum dissimilarity measure. The decision rule can be defined as Rule 1.

Rule 1: Classify the input shape as the kth reference shape, if $D_k = \min_{i=1}^{c}\{D_i\}$.

The proposed recognition algorithm can be summarized as follows.

1) Perform dominant point detection by the curvature-based polygonal approximation method for each reference shape to obtain the vertices of the approximated polygon [10].
2) Repeat steps 3–7 until there is no input shape to be recognized.
3) Perform dominant point detection on the input shape [10].
4) Find the global matching cost and the best-matched pair between the input shape and each of the reference shapes by the global cyclic string matching technique.
5) Compute the local dissimilarity measure between the input shape and each of the reference shapes by using the best-matched pair information.
6) Assess the dissimilarity measure by combining the global matching cost and the local dissimilarity measure.
7) Classify the input shape by Rule 1.

III. Experimental Results

In order to combine the reciprocal of the compactnesses, the lengths, and the angles in dissimilarity measure computing, it is necessary to make them dimensionless. The simple way of normalizing them is to divide them by their corresponding maximum values. Since the features used in these two stages are TRS-invariant, the proposed recognition algorithm is independent of changing of the position, orientation, and scale.

As mentioned in Section II-C, several approaches can be used to define the dissimilarity measure function by combining the global matching cost and the local dissimilarity measure. In the experiments, five types of dissimilarity measure functions f_i in (4) were applied. They are:

1) f_1(LD, GD) = GD;
2) f_2(LD, GD) = LD;
3) f_3(LD, GD) = (LD + GD)/2;
4) f_4(LD, GD) = min(LD, GD);
5) f_5(LD, GD) = max(LD, GD).

The first type is a conventional one-stage string matching technique. It uses the global matching cost as the dissimilarity measure in the recognition. The second type is a two-stage string matching approach, which considers only the local dissimilarity measure in the recognition process. Types three to five combine both global and local information in defining the dissimilarity measures. The recognition rates of the cyclic string matching under these five types of dissimilarity measure computing will be compared in the experiment.

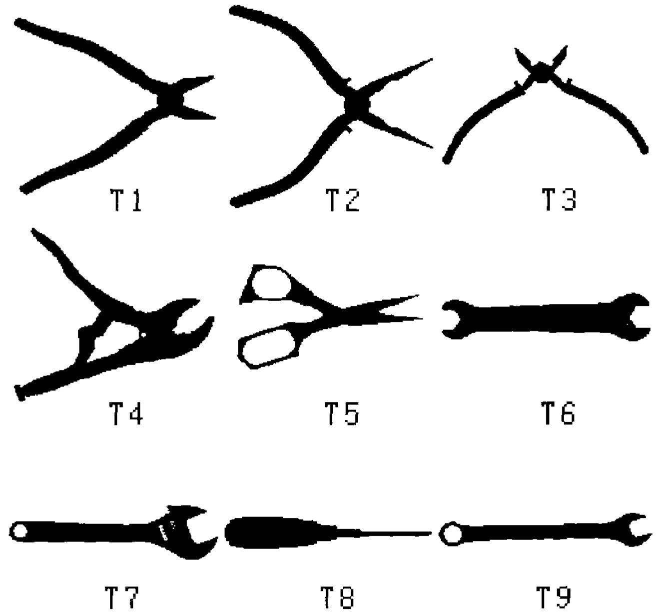


Fig. 6. Testing images of the nine hands tools.

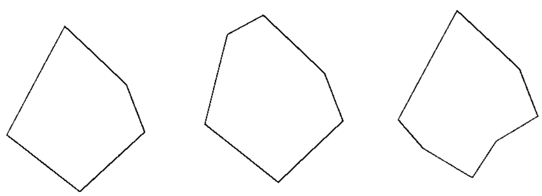
In this experiment, the testing images of nine different hand tools were used for evaluation (see Fig. 6). Since a good object recognition method should be robust for different orientations and scalings, for each tool image, there are 16 different orientations and four different scales conducted in the experiment. The 16 different orientations were arbitrary chosen by rotating the tools and the positions of the tools were changed at the same time. For each orientation, three additional images were generated by reducing the image to 90%, 80%, and 70% of original in both the x and y dimensions. Thus, 64 (16 × 4) testing images for each tool were used for recognition, and a total of 576 (9 × 64) testing images were used in the experiment. Besides, the opening levels of the tools 1–5 were fixed. The matching algorithm was applied to each testing image for recognition. If a wrong classification was made, an error was recorded. And the recognition rate can be computed. The experimental results are shown in Table I. The data in Table I are the recognition rates for the nine tools under five different measures.

For tool 7 in Table I, it is seen that the use of the conventional one-stage string matching has only 52% recognition rate. It is due to that tool 7 is very similar to tool 9. But when using the local dissimilarity measure or the combined dissimilarity measure, a significant improvement in recognition rate is found. Moreover, for tool 8, the one-stage string matching method has 84% recognition rate. But when using the local dissimilarity measure, the recognition rate is only 64%. It appears that high recognition rates are obtained only when averaging the global matching cost and the local dissimilarity measure. This demonstrates the advantages of using two-stage string matching.

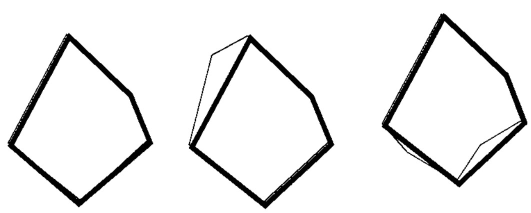
Furthermore, to examine the recognition rates displayed in Table I, It can be seen that when only the global matching cost, GD, is considered in the computing dissimilarity measure, we have a recognition rate of 88%. On the other hand, if we only use the local dissimilarity measure, LD, we have a recognition rate of 89%. Thus there is no significant improvement by conducting two-stage string matching if only the local dissimilarity measure is considered. However, if we combine the global matching cost and the local dissimilarity measure to obtain the dissimilarity measure,

TABLE I
COMPARISON OF THE RECOGNITION RATES (%) FOR THE NINE TOOLS UNDER FIVE DIFFERENT DISSIMILARITY MEASURES

TOOL	OVERALL DISSIMILARITY MEASURES				
	f_1	f_2	f_3	f_4	f_5
1	98	97	100	98	100
2	100	94	100	100	98
3	86	100	100	100	95
4	100	98	100	100	100
5	100	95	100	100	97
6	90	92	91	91	92
7	52	75	78	78	70
8	84	64	91	83	83
9	87	89	91	91	87
Average	88	89	95	93	91



(a)



(b)

Fig. 7. (a) Example of segmentation inconsistency problem and (b) the advantage of using the matched vertices rather than using the original vertices in the local dissimilarity measures computing.

the recognition rates reach to 95%, 93%, and 91% for f_3, f_4, and f_5, respectively. It is clear that combining the global matching cost and the local dissimilarity measure tends to have a better performance in recognition rate than that of using only the local dissimilarity measure.

IV. CONCLUSIONS

String matching is a useful tool for 2-D object recognition, but it tends to be affected by uneven segmentation problem. In this correspondence, we propose a new two-stage string matching technique for 2-D object recognition. The global string matching is conducted in the first stage. And a local dissimilarity measure is computed in the second stage. The global matching cost and the local dissimilarity measure are then combined to obtain the dissimilarity measure between the two objects. Finally, the input shape can be classified into one of the reference shapes by using the minimum dissimilarity measure.

In order to illustrate the merit of using the matched vertices in the local dissimilarity measures computing, an example is shown in Fig. 7. The three shapes in Fig. 7(a) are from the same object except that they have different edge detection results; thus they have different polygonal approximations. The matching costs of these three shapes will have large differences if the lengths and angles are used as attributed symbols. This is the segmentation problem as indicated by Maes [3].

The matching cost can be greatly varied due to the segmentation inconsistency, and it may not be able to really reflect the dissimilarity measure between two objects. But the best-matched pair can provide us the one-to-one matching information between the vertices of the two polygons. By using the best-matched pair information as shown in Fig. 7(b), the three shapes in Fig. 7(a) will have rather similar results in lengths and angles [3].

From the experimental results, it is seen that the two-stage string matching method has an improved recognition rate while comparing to the conventional one-stage string matching method. Another advantage of the proposed method is that it does not need to set any parameter in the recognition process.

REFERENCES

[1] K. S. Fu, "A step toward unification of syntactic and statistical pattern recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, pp. 200–205, 1983.

[2] M. Maes, "On a cyclic string-to-string correction problem," *Inform. Process. Lett.*, vol. 35, pp. 73–78, 1990.

[3] ——, "Polygonal shape recognition using string-matching techniques," *Pattern Recognit.*, vol. 24, pp. 433–440, 1991.

[4] S. Marshall, "Review of shape coding techniques," *Image Vis. Comput.*, vol. 7, pp. 281–294, 1989.

[5] D. Sankoff and J. B. Kruskal, Ed., *Time Warps, String Edits and Micromolecules: The Theory and Practice of Sequence Comparison.* Reading, MA: Addison-Wesley, 1983.

[6] W. H. Tsai and K. S. Fu, "Attributed grammar—A tool for combining syntactic and statistical approaches to pattern recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-10, pp. 873–885, 1988.

[7] W. H. Tsai and S. S. Yu, "Attributed string matching with merging for shape recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 7, pp. 453–461, 1985.

[8] Y. T. Tsay and W. H. Tsai, "Model-guided attributed string matching by split-and-merge for shape recognition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 3, pp. 159–179, 1989.

[9] R. A. Wagner and M. J. Fischer, "The string-to-string correction problem," *J. ACM*, vol. 21, pp. 168–173, 1974.

[10] W. Y. Wu and M. J. Wang, "Detecting the dominant points by the curvature-based polygonal approximation," *CVGIP: Graph. Models Image Process.*, vol. 55, pp. 79–88, 1993.

[11] ——, "Two-dimensional object recognition by using the reciprocal of compactnesses," in *Proc. 1994 Workshop on CVGIP*, 1994.

[12] K. C. You and K. S. Fu, "A syntactic approach to shape recognition using attributed grammars," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-9, pp. 334–345, 1979.