A Geometrical Approach to Polygonal Dissimilarity and the Classification of Closed Boundaries

R.L. Kashyap* and B.J. Oommen**

SCS-TR-18
January 1983

* Department of Electrical Engineering, Purdue University, W. Lafayette, IN, 47907, U.S.A.

** School of Computer Science, Carleton University, Ottawa, Ontario, K1S 5B6, Canada.

To be published in the IEEE Transactions on Pattern Analysis and Machine Intelligence. Details of the Issue of Publication are not known.

A GEOMETRICAL APPROACH TO POLYGONAL DISSIMILARITY AND THE CLASSIFICATION OF CLOSED BOUNDARIES†

R. L. Kashyap and B. J. Oommen

ABSTRACT

The problem of quantizing the dissimilarity between two irregular polygons has been considered. Two measures which are geometrical in nature and which capture the intuitive notion of the dissimilarity between shapes have been presented. Both these measures are related to the minimum value of the common area of the polygons when they are superposed on one another in various configurations. The first of these measures is edge based and the second is vertex based. Both are pseudo-metrics and are equally informative. An alternative measure of dissimilarity, referred to as the minimum integral square error between the polygons has also been proposed. The latter is closely related to the the area based dissimilarity measures, but is more easily computable. Using the minimum integral square error as a criterion, pattern classification of closed boundaries can be performed. Experimental results involving the boundaries of the four Great Lakes, Erie, Huron, Michigan and Superior justify the theoretical results presented in this paper.

† Partially supported by the National Science Foundation under grant ECS-80-09041.

Index Terms

Polygonal Dissimilarity, Minimum Integral Error Dissimilarity Measures, Area Dissimilarity Measures, Geometrical Pattern Classification, The Great Lakes.

I. INTRODUCTION

This paper considers the quantification of the dissimilarity between two irregular polygons, normalized to have either the same area or the same perimeter, and the use of the dissimilarity measure in the classification of a closed boundary into one of a finite number of classes. The theory of pattern classification of closed boundaries has a myriad of applications, including the automatic recogniton of characters[3,13,15,24], chromosomes[13], chairs[35], industrial parts[16, 31], keys[36], maps[7], and airplanes[8, 9, 25, 33, 34, 37]. With no loss of generality we assume that the boundaries are specified as irregular polygons.

Most of the current pattern classification procedures for the classification of closed boundaries implicitly or explicitly involve a distance measure for the comparison of two polygons. A careful study of these existing measures reveals some major defeciencies, primarily because none of these measures take into account all the geometrical features of the polygons. One can conjecture that a dissimilarity measure between polygons which emphasizes their geometrical properties would give better results in the pattern classification of closed boundaries than the measures currently in use.

We will briefly review the literature on distance measures between polygons and the related recognition problem. In the first approach, a set of global numerical features are extracted from each polygon. Some popular feature sets are the fourier descriptors and their variants[15, 24, 25, 33, 35, 38], moments[3, 9, 17], chord distributions[30], and the circular autoregressive model parameters[18]. Some of these feature sets possess interesting properties such as invariance to translation, rotation and scaling of the polygon. The distance between a pair of polygons is defined in terms of a norm in the feature space, for example, the Eucledian metric. Since both the above steps have only indirect relevance to the shapes of the boundaries, it is not too hard to see that the above measures can only partially capture the intuitive geometric concept of the differences in the shapes of two boundaries. In other words, it is possible to construct a pair of boundaries which are significantly different in shape and yet have a relatively small value of the dissimilarity measure which involves these features. However, our observation does not render the recognition algorithm using such descriptors useless, since there will be many situations where the defeciency mentioned above does not cause any major problem.

The second approach[5, 10, 11, 12, 13], is to reduce each polygon into a linear string of symbols. The distance between the two polygons is defined as the distance between their respective string representations, such as their correlation[10]. The Levenshtein metric[32] between the two strings can also be used after making due allowance for the fact that the strings represent <u>closed</u> figures. However the Levenshtein metric has not been used for the classification of boundaries, because it suffers from the drawbacks mentioned earlier in that it does not capture the intuitive notion of the dissimilarity between two shapes. Some of these distance measures are especially sensitive to noise[see the comments in 7]. Hence it is possible to have two boundaries which are similar in shape but which are quite distinct as viewed from their distance measure and vice versa. Similar comments are valid for the technique discussed in[13 (pp.61-62)].

There are many other methodologies for classification like the syntactic approach[1, 8, 13, 22, 23, 31, 36, 37] and the relaxation approach[7, 27] which do not explicitly need the concept of a dissimilarity measure between polygons. The syntactic techniques have been critically reviewed in the literature[22, 34]. Relaxation techniques have been used in the classification problem, but are usually slow and are computationally more expensive than both chain-code techniques and the techniques that utilize the global numerical properties of the polygons[7].

It is also appropriate to mention the existence of algorithms in artificial intelligence, which are used for testing whether two polygons are rotated, translated or scaled versions of one another[2, 29]. However, these algorithms do not deal with the dissimilarity in shape beteen two arbitary irregular polygons, and are thus of little use in pattern classification.

In this paper many geometric measures of dissimilarity between a pair of polygons are developed. The pair of polygons are suitably normalized to have either the same area or the same perimeter. The dissimilarity measures defined fall into two categories --- those which are edge based and those which are vertex based. An edge based based dissimilarity measure is defined by geometrically superposing the two polygons ξ and μ so that the ith edge of ξ falls alongside the jth edge of μ with their midpoints coinciding. Let $E_{i,j}(\xi,\mu)$ be the non-overlapping area between the polygons in this configuration, defined as the sum of the areas of the two polygons minus twice their overlapping area. The minimum of $E_{i,j}(\xi,\mu)$ over all i and j is defined as $E(\xi,\mu)$, the dissimilarity between the two polygons. The vertex based dissimilarity measure is analogously defined except that in the definition of their non-overlapping areas, $V_{i,j}(\xi,\mu)$, the polygons are superposed with the ith vertex of one coinciding with the jth vertex of the other with their angular bisectors falling alongside each other.

Such measures do not involve the arbitrary choice of features which always result in a loss of information[14]. The actual geometrical superposition ensures that the overall shape of the two polygons are considered in the dissimilarity measure, and not merely their global features. Further, both the edge based dissimilarity, $E(\xi,\mu)$ and the vertex based dissimilarity $V(\xi,\mu)$ between ξ and μ are zero if and only if ξ is a rotated, scaled or translated version of μ. It is intuitively plausible that both $E(\xi,\mu)$ and $V(\xi,\mu)$ are small if and only if the geometrical resemblance between ξ and μ is great.

One of the disadvantages of the non-overlapping area measure is that its evaluation is computationally time consuming. Here another geometrical measure is introduced which behaves like the non-overlapping area, but is easier to compute. This measure is termed as the minimum integral error between the polygons.

In this paper, algorithms will be developed for the classification of closed boundaries based on this dissimilarity measure. The given test boundary is first converted into an irregualr polygon. Usually this polygon has hundreds of edges. Since the computation of this dissimilarity measure is a monotonic function of the product of the number of edges of the polygons, it is imperative that the given polygon be smoothed to reduce the number of edges. Any of the smoothing techniques used in the literature[21] can be used for this purpose. In this study, a variant of the minimum perimeter polygon[19] has been used to smooth the polygons. Classification of the test boundary is acheived by computing the dissimilarity between its polygonal representation and the individual polygonal representations of the ideal boundaries of the various classes.

In Section II we present the various edge based and vertex based dissimilarity measures. In Section III we consider the computational aspects of the minimum integral error dissimilarity measure. In Section IV we propose the use of this measure in a pattern classification problem. We conclude the paper with the experimental results obtained in the study of the boundaries of the four Great Lakes, Erie, Huron, Michigan and Superior.

II. EDGE AND VERTEX BASED DISSIMILARITY

In this section we present two dissimilarity measures between two irregular polygons in terms of their areas and their intersecting areas. The polygons can have an arbitrary number of edges and their respective number of edges need not be equal. The first of these dissimilarity measures is edge based and the second is vertex based. We also propose two alternative measures, which are closely related to these but which are more easily computable.

Definition : A dissimilarity measure between two polygons ξ and μ is said to be a pseudo-metric if and only if :

(1) The dissimilarity measure between ξ and μ is exactly the dissimilarity between μ and ξ.

(2) The dissimilarity measure between ξ and μ is zero if and only if ξ and μ are scaled, rotated or translated versions of one another.

II.1. An Edge Based Area Dissimilarity (EBAD) Measure

Let ξ and μ be any two polygons normalized to either have the same area or the same perimeter, and let their areas be A_ξ and A_μ respectively. Let the polygons be superimposed with the ith edge of ξ falling alongside the jth edge of μ with the midpoints of these edges coinciding. Let $e_{i,j}(\xi,\mu)$ be the common area of the two polygons in this configuration. We define the dissimilarity between the polygons in this configuration as $E_{i,j}(\xi,\mu)$.

$$E_{i,j}(\xi,\mu) = A_\xi + A_\mu - 2e_{i,j}(\xi,\mu)$$

The Edge Based Area Dissimilarity (EBAD) measure between the polygons, $E(\xi,\mu)$, is defined as the minimum of $E_{i,j}(\xi,\mu)$ over all i and j. Thus,

$$E(\xi,\mu) = A_\xi + A_\mu - 2 \underset{i=1,..,N}{\text{Max.}} \quad \underset{j=1,..,M}{\text{Max.}} \left[e_{i,j}(\xi,\mu) \right] \qquad (2.1)$$

where ξ and μ have N and M edges respectively, N not necessarily equal to M.

Theorem I

The EBAD measure $E(\xi,\mu)$ between two arbitrary polygons ξ and μ is a pseudo-metric, i.e., $E(\xi,\mu)$ is symmetric and is equal to zero if and only if ξ is a rotated, scaled or translated version of μ.

Proof.

The symmetry of $E(\cdot,\cdot)$ follows by definition. Further, by observation, $E(\xi,\mu)=0$ if ξ is a rotated, scaled or translated verion of μ.

To prove the converse we note that $E(\xi,\mu)=0$ implies,

$$A_\xi + A_\mu = 2 \underset{i=1,..,N}{\text{Max.}} \quad \underset{j=1,..,M}{\text{Max.}} \left[e_{i,j}(\xi,\mu) \right]$$

Let the maximum value of $e_{i,j}(\xi,\mu)$ occur when i=p and j=q. Then,

$$A_\xi - e_{p,q}(\xi,\mu) = e_{p,q}(\xi,\mu) - A_\mu$$

The left hand side is a nonnegative quantity since $e_{p,q}(\xi,\mu)$ is an area which is always less than or equal to A_ξ. Similarly, the right hand side is a nonpositive quantity. Both these conditions can be satisfied if and only if $A_\xi=e_{p,q}(\xi,\mu)=A_\mu$, implying that the polygons exactly register on one another. Hence the theorem.

***

Besides being a pseudo-metric, the EBAD measure captures the intuitive geometrical aspects of distinguishing between shapes. Thus we expect the quantity $E(\xi,\mu)$ to be small if ξ and μ are intuitvely similar and the quantity to be relatively larger if the polygons are more dissimilar. We illustrate the EBAD measure by considering two simple geometrical figures.

Example I

Let ξ be the square ABCD and μ the irregular pentagon PQRST, normalized to have the same perimeter, as shown in Fig.II.1. Let the length of the edge of the square be unity, and the lengths of PQ and ST be a. Then, the length of QR is 1.5−a. By geometrical superposition we observe that the quantity $E(\xi,\mu)$ is identical to $E_{4,5}(\xi,\mu)$, which is the dissimilarity between the figures in the configuration shown in Fig.II.1. Using geometric and trigonometric arguments we obtain the following:

$$A_\xi = 1 \quad \text{and} \quad A_\mu = a + 0.5\sqrt{a^2 + 2 - 3a}$$

$$e_{4,5}(\xi,\mu) = 1 - \frac{0.5(1-a)^2}{\sqrt{a^2 + 2 - 3a}}$$

Hence, $E(\xi,\mu) = A_\xi + A_\mu - 2e_{4,5}(\xi,\mu)$

$$= a + \frac{\sqrt{a^2 + 2 - 3a}}{2} + \frac{(1-a)^2}{\sqrt{a^2 + 2 - 3a}} - 1.$$

This dissimilarity is a monotonically decreasing function of a in the interval [0,1], and it has the value of zero only when a is unity. This is the case when the pentagon reduces to a square. The dissimilarity is maximum when a is zero, which is a measure of the dissimilarity between the square and the corresponding isosceles triangle.

***

II.2. An Vertex Based Area Dissimilarity (VBAD) measure

Analogous to the edge based area dissimilarity measure between two polygons, we can define the Vertex Based Area Dissimilarity (VBAD) measure. As before, let ξ and μ be any two polygons normalized to either have the same area or the same perimeter. Let their areas be A_ξ and A_μ respectively. We now superimpose ξ and μ so that the ith vertex of ξ coincides with the jth vertex of μ, and their angular bisectors at these vertices fall alongside each other. Let $v_{i,j}(\xi,\mu)$ be the area of the intersection of the polygons in this configuration. The dissimilarity between the polygons in this configuration is given by $V_{i,j}(\xi,\mu)$.

$$V_{i,j}(\xi,\mu) = A_\xi + A_\mu - 2v_{i,j}(\xi,\mu)$$

The minimum of $V_{i,j}(\xi,\mu)$ over all i and j is defined as the Vertex Based Area Dissimilarity (VBAD) measure between ξ and μ. The expression for this measure, $V(\xi,\mu)$, is:

$$V(\xi,\mu) = A_\xi + A_\mu - 2 \underset{i=1,..,N}{\text{Max.}} \quad \underset{j=1,..,M}{\text{Max.}} \left[v_{i,j}(\xi,\mu) \right] \qquad (2.2)$$

where ξ and μ have N and M edges respectively, N not necessarily equal to M.

Theorem II

The VBAD measure between ξ and μ, given by $V(\xi,\mu)$ is a pseudo-metric.

***

The proof of the theorem is similar to the proof of Theorem I. The properties of the VBAD measure are illustrated using two simple geometrical polygons in Example II.

Example II

Let ξ be the square ABCD and μ the rhombus PQRS. Let the lengths of all their edges be unity, so that the two figures have the same perimeter. Let the minor angle of the rhombus be ψ. Then, $A_\xi=1$ and $A_\mu=\text{Sin}\psi$.

To compute $V(\xi,\mu)$ we need to consider only two cases, namely case (a), when the acute angle of the rhombus falls on the right angle of the square, and case (b), when the obtuse angle of the rhombus falls on the right angle of the square. These two cases are given in Fig.II.2 and Fig.II.3 respectively.

Consider Case (a). Using extensive geometric and trigonometric arguments it can be shown that the intersecting area PQKCSP is

$$1 - \mathrm{Sin}\frac{90-\psi}{2} - \frac{\mathrm{Sin}\frac{90-\psi}{2}\ \mathrm{Sin}\frac{450+3\psi}{4}\ \mathrm{Sin}\frac{90-\psi}{4}}{\mathrm{Sin}^2\frac{270+\psi}{4}}$$

Using A_ξ and A_μ, the value of the dissimilarity in this configuration simplifies to (2.3).

$$\mathrm{Sin}\psi + 2\mathrm{Sin}\frac{90-\psi}{2}\left[1 + \frac{\mathrm{Sin}\frac{450+3\psi}{4}\ \mathrm{Sin}\frac{90-\psi}{4}}{\mathrm{Sin}^2\frac{270+\psi}{4}}\right] - 1 \tag{2.3}$$

In Case (b), similar agruments yield the dissimilarity as given by (2.4).

$$1 + \mathrm{Sin}\psi\left[\frac{2\ \mathrm{Sin}\frac{90-\psi}{2}}{\mathrm{Sin}\frac{270-\psi}{2}} - 1\right] \tag{2.4}$$

The quantity $V(\xi,\mu)$ is the minimum of (2.3) and (2.4). A plot of these two expressions reveals that for all ψ the expression (2.3) is less than or equal to the expression (2.4). Hence $V(\xi,\mu)$ assumes the form of (2.3).

Note that $V(\xi,\mu)$ is a monotonically decreasing function of ψ having the value zero only when ψ=90, i.e., when the rhombus reduces to a square. The maximum dissimilarity is when ψ tends towards zero, in which case the rhombus tends towards becoming a straight line equally inclined to both the axes.

***

Although the quantities $E(\xi,\mu)$ and $V(\xi,\mu)$ quantify the dissimilarity between the polygons ξ and μ geometrically, they are not too easily computable. We shall present some other geometrical dissimilarity measures, which are more easily obtainable, and which are geometrically closely related to these above measures.

II.3. Integral Error Dissimilarity Measures

Let ξ and μ be two polygons normalized to have the same unity perimeter. We superimpose the two polygons so that the ith edge of ξ falls alongside the jth edge of μ and the midpoints of the two edges coincide. This midpoint can be made the origin of a new rectangular coordinate system, with the abscissa falling alongside these common edges. Let both the polygons be traversed in a clockwise direction starting from the origin. Let $H_i(\lambda)$ be the unique point reached on ξ after a traversal of length λ, and let $G_j(\lambda)$ be the unique point reached on μ after a traversal of λ, $\lambda \in [0,1]$. Any well defined norm in R^2 can now be used to quantize the pointwise dissimilarity between the points $H_i(\lambda)$ and $G_j(\lambda)$. In particular, we use the square error criterion and express this dissimilarity as $||H_i(\lambda) - G_j(\lambda)||^2$. The cumulative effect of this local dissimilarity is obtained by integrating the latter quantity over λ in the interval $[0,1]$. We refer to this integral as the integral square error between the two polygons relative to their i-j edges, written as $D_{i,j}(\xi,\mu)$. Symbolically,

$$D_{i,j}(\xi,\mu) = \int_0^1 || H_i(\lambda) - G_j(\lambda)||^2 \, d\lambda$$

The minimum of this quantity over all possible values of i and j is defined as the minimum integral square error, $D(\xi,\mu)$ between them.

$$D(\xi,\mu) = \underset{i=1,..N}{\text{Min.}} \quad \underset{j=1,..M}{\text{Min.}} \left[\int_0^1 || H_i(\lambda) - G_j(\lambda)||^2 \, d\lambda \right]$$

where ξ and μ have N and M edges respectively, N not necessarily equal to M.

<u>Theorem III</u>

The edge based minimum integral square error between ξ and μ given by $D(\xi,\mu)$ is a pseudo-metric.

***

The proof of the theorem is straightforward.

<u>Example III</u>

Let ABCDEF and PQRSTUW be two normalized polygons ξ and μ respectively, with unequal number of edges as in fig. II.4. The superposition indicated in the figure II.4 leads to the computation of $D_{1,1}$. The superposition indicated in the figure II.5 leads to the evaluation of $D_{1,5}$. By inspection, we can conclude $D=D_{1,1}$. In the figure II.6, we have graphed $||H_i(\lambda)-G_1(\lambda)||^2$ and $||H_1(\lambda)-G_5(\lambda)||^2$ versus λ.

The graph of $||H_i(\lambda)-G_j(\lambda)||^2$ versus λ for the optimal configuration (i,j) (i.e., when $D=D_{i,j}$) yields information about both local and global similarity of the 2 shapes. For 2 strongly similar shapes as in figure II.4, the quantity $||H_1(\lambda)-G_1(\lambda)||^2$ is uniformly small for all λ. On the otherhand, consider the 2 shapes in figure II.1 which share an identical common portion QPADTS. Here the optimal configuration is $(4,5)$. Since the graph of $||H_4(\lambda)-G_5(\lambda)||^2$ vs λ will be zero for $0 < (\lambda, 1-\lambda) < \lambda_1$, where $\lambda_1 > 1/2$ (length of the smaller of the 2 common edges), we can conclude that the 2 shapes have partial similarity.

We now proceed to obtain an explicit expression for the minimum integral square error between the two figures considered in Example II, and relate it to the edge based area dissimilarity measure $E(\cdot,\cdot)$.

Example IV

Consider the two figures ξ and μ given by the square ABCD, and the rhombus, PQRS, of Fig.II.7. Let ψ be the minor angle of the rhombus. Due to the symmetry of the figures, we need to consider only one configuration, for example, that of Fig.II.7, to compute $D(\xi,\mu)$. In this configuration, the origin, θ, is the midpoint of the common edges AD and PS. The value of λ corresponding to A, B, C, and D are 0.125, 0.375, 0.625 and 0.875 respectively, which are also the values of λ corresponding to P, Q, R, and S respectively. A sketch of the computation of $D(\xi,\mu)$ is given below.

$$\text{For} \quad 0\leq\lambda\leq 0.125, \quad \| H_1(\lambda)-G_1(\lambda)\|^2 = 0$$

$$\text{For } 0.125\leq\lambda\leq 0.375, \quad \| H_1(\lambda)-G_1(\lambda)\|^2 = 4(\lambda-0.125)^2 \cos^2 \frac{90+\psi}{2}$$

$$\text{For } 0.375\leq\lambda\leq 0.625, \quad \| H_1(\lambda)-G_1(\lambda)\|^2 = 0.25 \cos^2 \frac{90+\psi}{2}$$

$$\text{For } 0.125\leq\lambda\leq 0.375, \quad \| H_1(\lambda)-G_1(\lambda)\|^2 = 4(0.875-\lambda)^2 \cos^2 \frac{90+\psi}{2}$$

$$\text{For } 0.875\leq\lambda\leq 1.0, \quad \| H_1(\lambda)-G_1(\lambda)\|^2 = 0$$

$$\text{Hence, } D(\xi,\mu) = \cos^2 \frac{90+\psi}{2} \left[4 \int_0^{0.25} p^2\, dp + 0.25 \int_{0.375}^{0.625} d\lambda + 4 \int_0^{0.25} p^2\, dp \right]$$

$$= \frac{10}{192} (1 - \sin \psi).$$

Consider now the edge based area dissimilarity measure $E(\xi,\mu)$. Using (2.1), this can be shown to have the value,

$$E(\xi,\mu) = 0.0625 [1 - \sin\psi + \sin\psi \cos\psi]$$

If $1-\sin\psi = u$, then, $1 - \sin \psi + \sin\psi \cos\psi$ is given by $u+(1-u)\sqrt{2u-u^2}$, which is a monotonically increasing function of u in the interval [0,1]. Hence the quantity $D(\xi,\mu)$ is a monotonically increasing function of $E(\xi,\mu)$. This relationship confirms use of $D(\xi,\mu)$ as a measure of the dissimilarity between ξ and μ.

***

Remarks

(1) In the definition of $D_{i,j}(\xi,\mu)$, the square of the L_2 norm was used to measure the pointwise dissimilarity between $H_i(\lambda)$ and $G_j(\lambda)$. Any monotonically increasing function of a well-defined norm in R^2 can be used to obtain measures which are equally informative.

(2) By studying the geometry of the situation, we conjecture that for two arbitary polygons ξ and μ, the integral square error $D(\xi,\mu)$ is a monotonically increasing function of their edge based area dissimilarity measure $E(\xi,\mu)$.

(3) Just as we have defined the minimum integral error by making the edges of the polygons fall alongside each other, we can analogously work with the vertices coinciding with each other. The latter would yield a measure closely related to $V(\xi,\mu)$ defined as in (2.2). Intuitively we anticipate that the integral error dissimilarity defined in terms of the vertices will be just as informative as the integral error defined in terms of the edges. We choose to work with the latter formulation primarily for computational simplicity.

III. COMPUTATION OF $D(\xi,\mu)$

Obtaining an explicit expression for $D(\xi,\mu)$ involving two arbitrary polygons ξ and μ having different number of edges would be rather tedious and often impractical. In such cases, the measure $D(\xi,\mu)$ can be computed numerically by computing $D_{i,j}(\xi,\mu)$ for various values of i and j.

To compute $D_{i,j}(\xi,\mu)$ for any two indices i and j, the functions $H_i(\lambda)$ and $G_j(\lambda)$ are evaluated at K points, where the integer K is determined by Δ the step size of λ, and is independent of the number of edges in both ξ and μ. The integral

$$\int_0^1 || H_i(\lambda) - G_j(\lambda)||^2 \, d\lambda$$

is approximated by the sum,

$$\sum_{k=0}^{K-1} || H_i(\lambda_k) - G_j(\lambda_k)||^2 \, \Delta \quad , \quad \text{where } K\Delta = 1.$$

Since the above expression involves the computation of K values of $H_i(\lambda)$ and $G_j(\lambda)$, K squared errors, and 2K additions, the time required to compute any one integral is approximately a linear function of K.

If ξ and μ have N amd M edges respectively, the brute force method of evaluating $D(\xi,\mu)$ requires the computation of all the MN integrals $D_{i,j}(\xi,\mu)$. However, by computing $D_{i,j}(\xi,\mu)$ for a <u>subset</u> of the MN integrals, the same minimum can be effectively computed.

Let O_ξ and O_μ be the centers of gravity of ξ and μ respectively. Let $\delta^{i,j}$ be the distance between O_ξ and O_μ in the configuration when the ith edge of ξ falls alongside the jth edge of μ. From the geometry of the situation, we observe that if $D_{i,j}(\xi,\mu)$ attains its minimum when i=p and j=q, then $\delta^{p,q}$ will be relatively small. Conversely, if $\delta^{r,s}$ is relatively large, it is unlikely that $D_{r,s}(\xi,\mu)$ will be the minimum integral (see Fig.III.1). Using the quantities $\delta^{i,j}$ a proper subset Γ of the set $\{1,\ldots,N\} \times \{1,\ldots,M\}$ can be chosen so that we need to compute $D_{i,j}(\xi,\mu)$ only for those configurations $(i,j)\epsilon\Gamma$ for evaluating $D(\xi,\mu)$.

<u>III.1. The Set of Configurations to be Considered</u>

The pair (i,j) is said to be a <u>competing pair</u> if $\delta^{i,j}$ is less than or equal to a predefined threshold, say, δ^*. The set of all the competing pairs is defined as Γ, the <u>competing set</u>.

$$\Gamma = \{ (i,j) \mid (i,j) \in \{1,..,N\} \times \{1,..,M\}, \delta^{i,j} \leq \delta^* \}$$

By studying the particular contours under consideration, usually a <u>small</u> value of δ^* can be specified so that

$$D(\xi,\mu) = \underset{(i,j)\in\Gamma}{\text{Min.}} \left[D_{i,j}(\xi,\mu) \right] \tag{3.1}$$

It must be noted that the positions of O_ξ and O_μ relative to the vertices of ξ and μ respectively, need be computed <u>only once</u>. They are conveniently stored in their polar representations. Subsequently, to test if any pair (i,j) is in Γ, the locations of O_ξ and O_μ in this configuration can be computed using <u>merely</u> one linear transformation.

The question of obtaining a suitable value for δ^* must now be considered. In a practical pattern recognition system, we are usually given training samples for each class. For each training sample, the value of $\delta^{i,j}$ for the optimum integral can be obtained for the various classes. Using these, a conservative estimate for δ^* can be obtained for each class. Occasionally, the value of δ^* may be underestimated, but this is not too serious, since the expression defined in (3.1) will then have a value close to $D(\xi,\mu)$, and will thus be a good approximation to it.

An alternative approach is to order the pairs (i,j) in the set $\{1,..,N\} \times \{1,..,M\}$ in ascending values of $\delta^{i,j}$. The integral $D_{i,j}(\xi,\mu)$ is computed for the first Z pairs. A safe value for the number of elements in Γ can be estimated from the polygons under consideration. A good rule of thumb which yields a conservative set Γ is to make $\#\Gamma=M+N$. In such a case, out of the MN distinct integrals, only M+N of them will even be attempted.

An upper bound for the complexity of the computation of $D(\xi,\mu)$ is obtained as below. Let a_1 be the cost associated with testing whether a pair (i,j) is in Γ, and let a_2 be the cost associated with evaluating any one integral $D_{i,j}(\xi,\mu)$. Then the total time required to compute $D(\xi,\mu)$ would be of the form $a_1MN+a_2\#\Gamma$. Since the cost for the former computation is usually negligible compared to the cost associated with the computation of an integral, the total computation time is approximately linear in $\#\Gamma$, the number of elements in Γ, the set of competing pairs.

III.2. Additional Reduction in Computation

In practice, all the integrals corresponding to the pairs in Γ need not be completely evaluated, since we are interested only in the one that yields the minimum. Initially, $D(\xi,\mu)$ is assigned the value of the first integral computed. If in the subsequent computation of any $D_{i,j}(\xi,\mu)$, the value of the integral exceeds the currently stored value of $D(\xi,\mu)$, the computation of that integral is terminated. If however the integral is evaluated to its completion, the value of $D(\xi,\mu)$ is updated as the minimum of its current value and the latest $D_{i,j}(\xi,\mu)$. This reduces the compuation time considerably.

In a typical example involving the boundaries of the Great Lakes Erie and Michigan, their polygonal approximations had 17 and 21 edges respectively. Using a conservative δ^* of value approximately 0.05, the set of competing pairs had <u>only</u> 27 elements. Of the 27 integrals which were attempted, only 3 of them had to be completely evaluated. In all the remaining 24 cases, the computation was terminated before the integral was evaluated completely, because its current value exceeded the most updated value of the dissimilarity measure. In this case the time required to compute $D(\xi,\mu)$ on the PDP 11/70 UNIX system was only 3.2 seconds.

IV. CLASSIFICATION USING POLYGONAL DISSIMILARITY

Consider a pattern clasification problem with J classes. Let ξ_i be the ideal polygon associated with the ith class. Let μ' be the test boundary to be classified, approximated by the polygon μ. The classification of μ can be performed by computing $D(\xi_i,\mu)$ for $i=1,\ldots,J$ and the pattern assigned to the class which minimizes this dissimilarity measure.

Since the ξ_i's and μ usually have a large number of edges, the computation of $D(\xi_i,\mu)$ for any i is time consuming. Hence it is better to smooth the polygons under consideration, before evaluating the dissimilarity measure. The smoothed version of any polygon τ, is a polygon τ^* which approximates it according to some criterion and which possesses fewer edges. The actual smoothing process can be performed by using any of the techniques known in the literature[21], such as the split and merge technique. We prefer to smooth the polygons by using a variant of its minimum perimeter polygon[19]. This variant, termed as its Linear Minimum Perimeter Polygon (LMPP), has some interesting scale preserving properties and is described in the next subsection. The actual classification is acheived by computing the dissimilarities between the smoothed versions of the ξ_i's and the smoothed version of μ.

If the number of classes is large, the classification can be acheived using a pyramidal approach. Initially, the ideal patterns of the various classes and the test pattern are crudely approximated, so that the number of edges in their smoothed versions is relatively small. The dissimilarities between the smoothed polygons are computed for all the J classes. The classes are now ordered in ascending orders of their dissimilarities, and a subset of the classes which have the smallest dissimilarities are chosen for a more precise examination. Typically, this subset may contain about 20% of the number of classes in the original problem. Finer approximations of the ideal boundaries and the test boundary are now used to more precisely assign the test boundary to one of the classes in this subset.

IV.1 Smoothing by the Linear Minimum Perimeter Polygon (LMPP) Approach

Let τ be any polygon whose smoothed version τ^* is to be obtained. The goal of the LMPP, τ^* is to merge the adjacent edges in τ which are almost collinear. The extent of smoothing is controlled by a parameter ε, $0 \leq \varepsilon < 1$. The larger the value of ε the smaller will be the number of edges in τ^*.

We precisely define the LMPP τ^* of a polygon τ as below. Let τ be defined by the vertices P_i', $i=1,\ldots,R$. Let τ_1 be any smoothed version of τ described by the ordered set $\{P_i, i=1,\ldots,R\}$. The polygon τ^*, which is obtained by the following minimization procedure is termed as the LMPP of τ.

$$\tau^* = \text{Argument } \underset{\tau_1}{\text{Min.}} \left[f(\tau_1) \right]$$

$$\text{subject to} \quad \| P_i' - P_i \| \leq \varepsilon L \quad i=1,\ldots\ldots,R \tag{4.1}$$

where L is the perimeter of τ and,

$$f(\tau_1) = \sum_{i=1}^{R} \| P_i - P_{i+1} \| \quad \text{with } P_{R+1} = P_1$$

The formulation of the LMPP τ^* is similar to the formulation of the minimum perimeter polygon by Montanari[19] except that the right hand side of (4.1) is a linear function of the perimeter of τ.

By the theory of optimization, the polygon τ^* is completely described by those vertices at which the constraints given by (4.1) are equality constraints. Hence,

$$\tau^* = \{ P_{i_k}, k=1,\ldots,N\}$$

where $|| P_{i_k}' - P_{i_k} || = \varepsilon L, \quad k=1,\ldots N, \quad N \leq R.$

It can be proved that the LMPP τ^* has the following desirable properties if the constraint regions given by (4.1) are disjoint.

(1) τ^* is exactly τ if and only if $\varepsilon = 0$.

(2) Let τ and η be two polygons, where η is a scaled version of τ, the scaling factor being k>0. Then their LMPPs τ^* and η^* computed for the same ε are exactly scaled versions of one another, and the scaling factor is exactly k.

(3) If τ^* is normalized to have a perimeter of unity, the sequence of angles and sides of the normalized polygon is independent to scaling and translation of the coordinate axes.

Example VI

Let τ be the quadrilateral ABCD, given in Fig.IV.1. The radii of the constraint disks around these vertices are given by εL, where

$$L = || AB || + || BC || + || CD || + || DA ||.$$

The LMPP of τ is τ^* given by the triangle PQR.

***

V EXPERIMENTAL RESULTS

A pattern classification experiment was performed using the boundaries of the four Great Lakes, Erie, Huron, Michigan and Superior. The maps of the lakes were obtained from [20] drawn at a scale of 32 miles/inch. They were appropriately scaled to fit an 8" X 11" frame. These pictures were photographed using a television camera to fit a 90 X 90 pixel array. A simple boundary tracking algorithm utilizing constant thresholds was then employed to extract the boundaries of the lakes.

V.1. Generating Noisy Boundaries

A procedure similar to the above was used to obtain test boundaries for classification. To provide a difficult matching environment for the classification process, unsmoothed boundaries were noisily degraded as follows. Let P_i be any arbitary point on the boundary, whose adjacent vertices are P_{i-1} and P_{i+1}. Let the centroid of the triangle $P_{i-1}\ P_i\ P_{i+1}$ be the point Q_i. Then the point P_i is noisily displaced to the point R_i which is on the line joining P_i and Q_i, where

$$R_i = P_i + \lambda(Q_i - P_i)$$

and λ is a Gaussian random variable of mean zero and variance σ^2. When $0 \leq \lambda \leq 1$, the point R_i is a point in between P_i and Q_i. Whenever λ is negative or greater than unity, the point lies on the infinite line joining P_i and Q_i, but is not in between them. This noise generating mechanism ensured a global and a local deformation on the boundary. In an example, when σ^2 was 6.25, a typical value of λ can be 5. This corresponds to a perturbation of 500% of the point P_i relative to the local centroid Q_i. The noise level of the noisy boundaries generated by this mechanism is much larger than the noise level obtained by mere variations in thresholding. The ideal boundaries and some typical noisy boundaries of the four Great Lakes are given in Fig. V.1 with $\sigma^2 = 6.25$. Note that the distortions in the shape are significant.

V.2. Classification Results

The reference boundaries and the noisy boundaries were smoothed using the same ε value of 0.005 after appropriately processing them to ensure that the constraint disks defined by (4.1) are disjoint. The smoothed versions of the noisy boundaries were compared to the smoothed versions of the reference boundaries using the minimum integral square error measure, and the test boundary assigned to the class which minimized this dissimilarity.

100 noisy boundaries were tested using 20 different maps and 5 different variance levels ranging from 1.0 to 9.0. Out of the 100, 97 of them were correctly classified. The details of the classification are presented in Table V.1.

We also graphically demonstrate how the measure $D(\cdot,\cdot)$ is able to discriminate between two classes which are distinct. The value of D(Ideal Erie, μ) is plotted against the corresponding value of D(Ideal Huron, μ) for various noisy boundaries of both Lake Erie and Lake Huron. The plot is given in Fig.V.2. From the figure, one can observe, that just as as the two lakes are noticably dissimilar, the values of the dissimilarity measures are also widely different. The clustering of the noisy boundaries naturally seperates them into the two classes. These results justify the use of the geometrical dissimilarity measures for the classification of closed boundaries.

VI. CONCLUSIONS

In this paper we have considered the problem of quantizing the dissimilarity between two irregular polygons ξ and μ. Two geometrical measures, namely the edge based area dissimilarity measure, $E(\xi,\mu)$, and the vertex based dissimilarity measure, $V(\xi,\mu)$, have been proposed. Both these measures are pseudo-metrics and appear to represent the intuitive concept of dissimilarity in shapes. A third measure, $D(\xi,\mu)$, termed as the minimum integral square error between ξ and μ has been proposed. The latter measure is closely related to $E(\xi,\mu)$, but is much more easily computable. The computation of $D(\xi,\mu)$ has been presented, and its use in pattern classification has been verified in experiments involving the four Great Lakes Erie, Huron, Michigan and Superior.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation to Professor Douglas McCarthy, of the Department of Computer Science at Purdue University, W. Lafayette, and to Mr. Bruce Nearing, a Purdue graduate presently with The Boeing International, for their invaluable contributions toward this study.

REFERENCES

1. Aho, A. V., and Peterson, T. G., "A Minimum Distance Error-Correcting Parser for Context-Free Languages", SIAM J. Comput. Vol.1, 1972, pp.305-312.

2. Akl, S. G., "An Improved Algorithm to Check for Polygonal Similarity", Proc. of the IEEE Comp. Soc. Confrence on Pat. Recog. and Image Process., 1978, pp.39-41.

3 Alt, F. L., "Digital Pattern Recognition by Moments", J. ACM. Vol.11, 1962, pp. 240-258.

4. Attneave, F. "Some Informational Aspects of Visual Perception", Psychological Review, Vol.61, 1954, pp.183-193.

5. Bribiesca, E., and Guzman, A. "How to Describe Pure Form, and How to Measure Differences in Shape Using Shape Numbers", Pattern Recognition, Vol. 12, 1980, pp.101-112.

6. Davis, L. S., "Understanding Shape: Angles and Sides", IEEE Trans. on Computers, Vol.C-26, 1977, pp.236-242.

7. Davis, L. S., "Shape Matching Using Relaxation Techniques", IEEE Trans. on Pat. Anal. and Mach. Intel. Vol.PAMI-1, 1979, pp.60-72.

8. Davis, L.S., and Henderson, T. C., "Heirarchial Constraint Processes for Shape Analysis", IEEE Trans. on Pat. Anal. and Mach. Intel. Vol.PAMI-3, 1981, pp.265-277.

9. Dudani, S. A., Breeding, K. J., and McGhee, R. B., "Aircraft Identification by Moment Invariants", IEEE Trans. on Computers, Vol.C-26, 1977, pp.39-45.

10. Feder, J., and Freeman, H., "Segment Fitting of Curves in Pattern Analysis Using Chain Correlation", AD619525, March 1965.

11. Freeman, H., "On the Encoding of Arbitrary Geometric Configurations", IRE Trans. on Computers, Vol.EC-10, 1961, pp.260-268.

12. Freeman, H., "Computer Processing of Line-Drawing Images", Computing Surveys, Vol.6, 1974, pp.57-97.

13. Fu, K. S., "Syntactic Methods in Pattern Recognition", Academic Press, New York/London, 1974.

14. Fu, K. S., "Sequential Methods in Pattern Recognition and Machine Learning", Academic Press, New York, 1968.

15. Grandlund, G. H., "Fourier Preprocessing for Hand Print Character Recognition", IEEE Trans. on Computers, Vol.C-21, 1972, pp.195-201.

16. Horaud, R., and Charras, J. P., "Automatic Inspection and Orientation of External Screws", Proceedings of the IEEE Int. Conf. on Pat. Recog. 1980, pp.264-268.

17. Hu, M-K., "Visual Pattern Recognition by Moment Invariants", IRE Trans. on Information Theory, Vol.IT-8, 1962, pp.179-187.

18. Kashyap, R. L., and Chellappa, R., "Stochastic Models for the Analysis of Closed Boundaries: Representation and Reconstruction", To appear in IEEE Trans. on Info. Theory. Vol.IT-27, 1981.

19. Montanari, U., "A Note on Minimal Length Polygonal Approximation to a Digitized Contour", Comm. of the ACM, Vol.13, 1970, pp.41-47.

20. The National Geographic Magazine. Maps of the Great Lakes Region. Drawn Dec. 1953.

21. Pavlidis, T., "Structural Pattern Recognition", New York: Springer-Verlag, 1977.

22. Pavlidis, T., "Syntactic Pattern Recognition of Shape", Proc. of the IEEE Comp. Soc. Conf. on Pat. Recog. and Image Processing, June 1977, pp.98-107.

23. Pavlidis, T., and Ali, F., "A Heirarchial Syntactic Shape Analyser", IEEE Trans. on Pat. Anal. and Mach. Intel. Vol. PAMI-1, 1979, pp.2-9.

24. Persoon, E., and Fu, K.S., "Shape Discrimination using Fourier Descriptors", IEEE Trans. on Systems, Man and Cybern. Vol. SMC-7, 1977, pp.170-179.

25. Richard, C.W., and Hemami, H., "Identification of Three-Dimensional Objects Using Fourier Descriptors of the Boundary Curve", IEEE Trans. on Systems, Man and Cybern. Vol. SMC-4, 1974, pp.371-378.

26. Rosenfeld, A., "Survey, Picure Processing : 1978", Computer Graphics and Image Processing, Vol.9, 1979, pp.354-393.

27. Rosenfeld, A., Hummel, R.A., and Zucker, S.W., "Scene Labelling by Relaxation Operations", IEEE Trans. on Systems, Man and Cybern. Vol. SMC-6, 1976, pp.420-433.

28. Sakane, S., "An Algorithm for Scale Invariant Segment Matching", Proc. of the IEEE Conf. on Pat. Recog. and Image Proc., 1981, pp.565-571.

29. Shamos, M.I., "Geometrical Complexity", Proc. of the Seventh Annual ACM Symposium on the Theory of Computing, Albuquerque, 1975, pp.224-233.

30. Smith, S.P., and Jain A.K., "Chord Distributions for Shape Matching", Proc. of the IEEE Conf. on Pat. Recog. and Image Proc., 1981, pp.168-170.

31. Tsai, W.H., and Fu, K.S., "A Syntactic-Statictical Approach to Recognition of Industrial Objects", Proc. of the Int. Conf. on Pat. Recog., 1980, pp.251-259.

32. Wagner, R.A., and Fischer, M.J., "The String to String Correction Problem", J. ACM, Vol.21, 1974, pp.168-173.

33. Wallace, T.P., and Wintz, P.A., "An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors", Computer Graphics and Image Processing, Vol.13, 1980, pp.99-126.

34. Wallace, T.P., Mitchell, O.R., and Fukunaga, K., "Three-Dimensional Shape Analysis using Local Shape Descriptors", IEEE Trans. on Pat. Anal. and Mach. Intel. Vol.PAMI-3, 1981, pp.310-323.

35. Watson, L.T., and Shapiro, L.G., "Identification of Space Curves from Two-Dimensional Perspective Views", Tech. Rep. CS 80004-R, Dept. of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA : 24061.

36. Yang, Y.H., and Sze, T.W., "A Simple Contour Matching Algorithm", Proc. of the IEEE Conf. on Pat. Recog. and Image Proc., 1981, pp.562-564.

37. You, K.C., and Fu, K.S., "Distorted Shape Recognition Using Attributed Grammars and Error-Correcting Techniques", Computer Graphics and Image Processing, Vol.13, 1980, pp.1-16.

38. Zahn, C.T., and Roskies, R.Z., "Fourier Descriptors for Plane Closed Curves", IEEE Trans. on Computers, Vol.C-21, 1972, pp.269-281.

Lake	Variance	Number of Boundaries	Number Correct	Misclassified Class
Erie	1.00	5	5	---
	2.25	5	5	---
	4.00	5	5	---
	6.25	5	5	---
	9.00	5	5	---
Huron	1.00	5	5	---
	2.25	5	5	---
	4.00	5	5	---
	6.25	5	5	---
	9.00	5	5	---
Michigan	1.00	5	5	---
	2.25	5	3	Superior
	4.00	5	5	---
	6.25	5	5	---
	9.00	5	5	---
Superior	1.00	5	5	---
	2.25	5	5	---
	4.00	5	5	---
	6.25	5	4	Michigan
	9.00	5	5	---

Total Classification Accuracy = 97%.

Table V.1. : The details of the pattern classification experiment performed involving the boundaries of the four Great Lakes, Erie, Huron, Michigan and Superior.

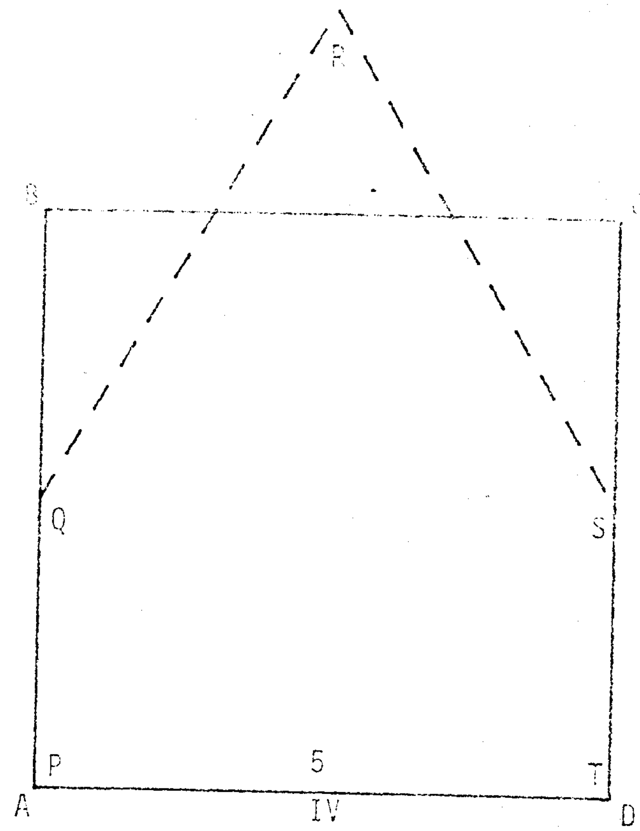

Fig. II.1. Computation of $E_{4,5}(\xi,\mu)$ between the square ξ, ABCD, and the irregular pentagon μ, PQRST. DA is the fourth edge of ξ, and TP is the fifth edge of μ.

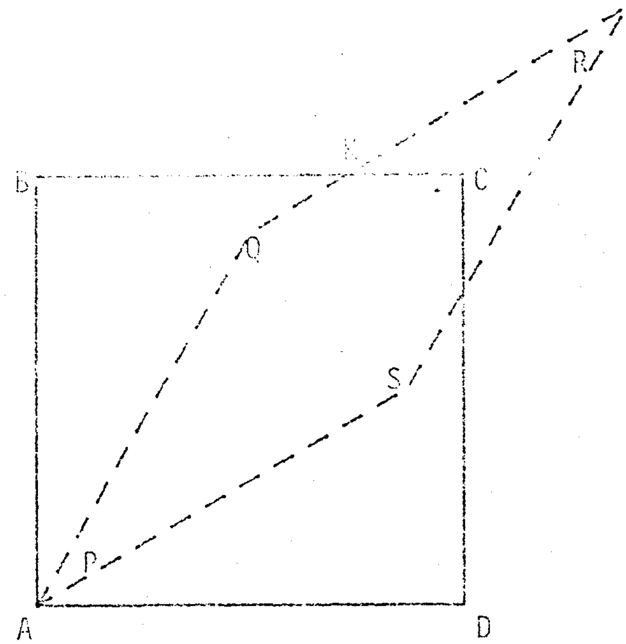

Fig. II.2. Case (a) of Example II for the computation of $V(\xi,\mu)$ between the square ξ and the rhombus μ.

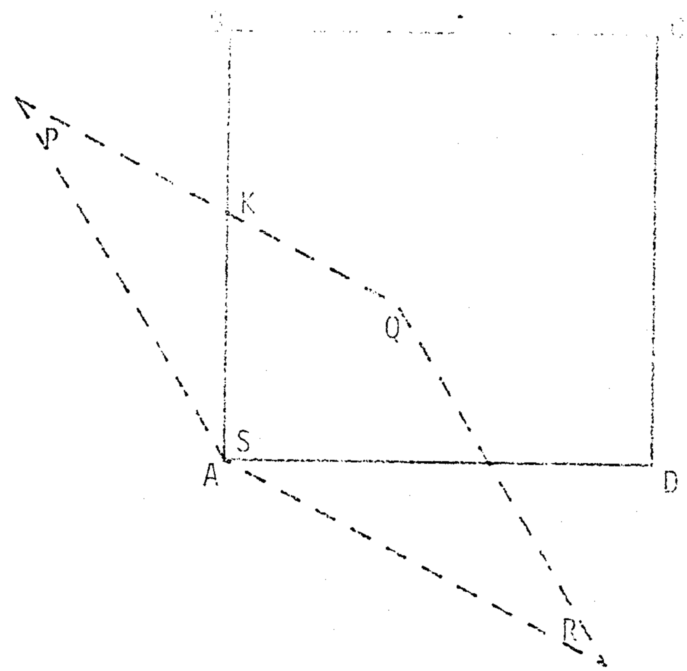

Fig. II.3. Case (b) of Example II for the computation of $V(\xi,\mu)$ between the square ξ and the rhombus μ.

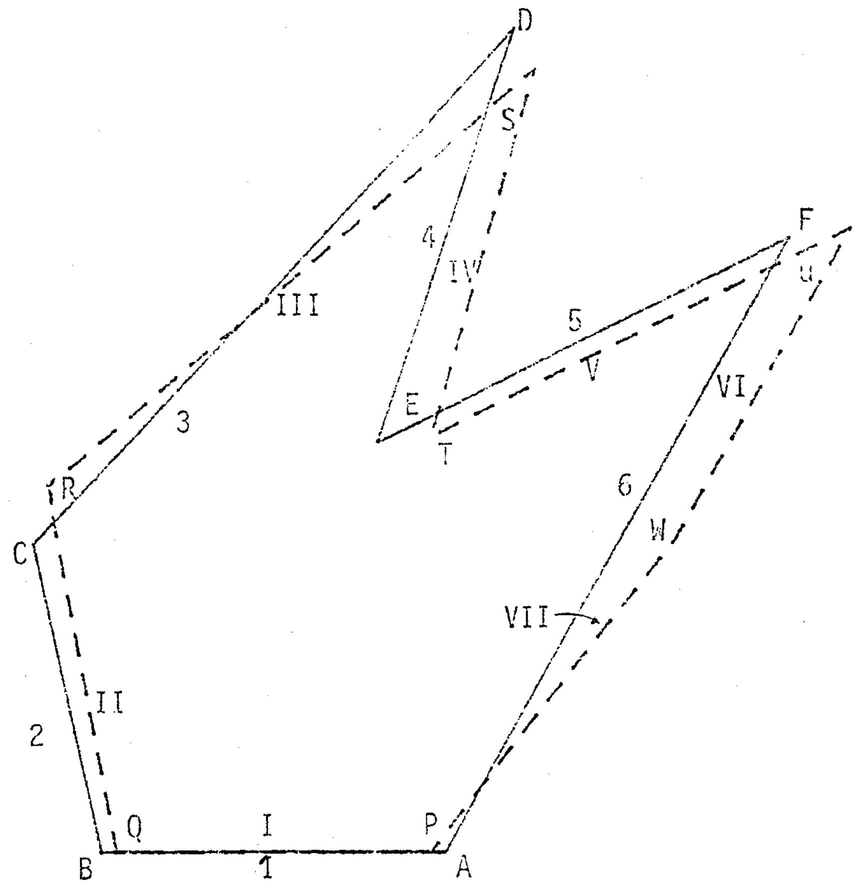

Fig. II.4. Computation of $D_{1,1}(\xi,\mu)$ between the hexagon ξ, ABCDEF, and the heptagon μ, PQRSTUW. The edges of ξ and μ are given by the Arabic and Roman numberals respectively.

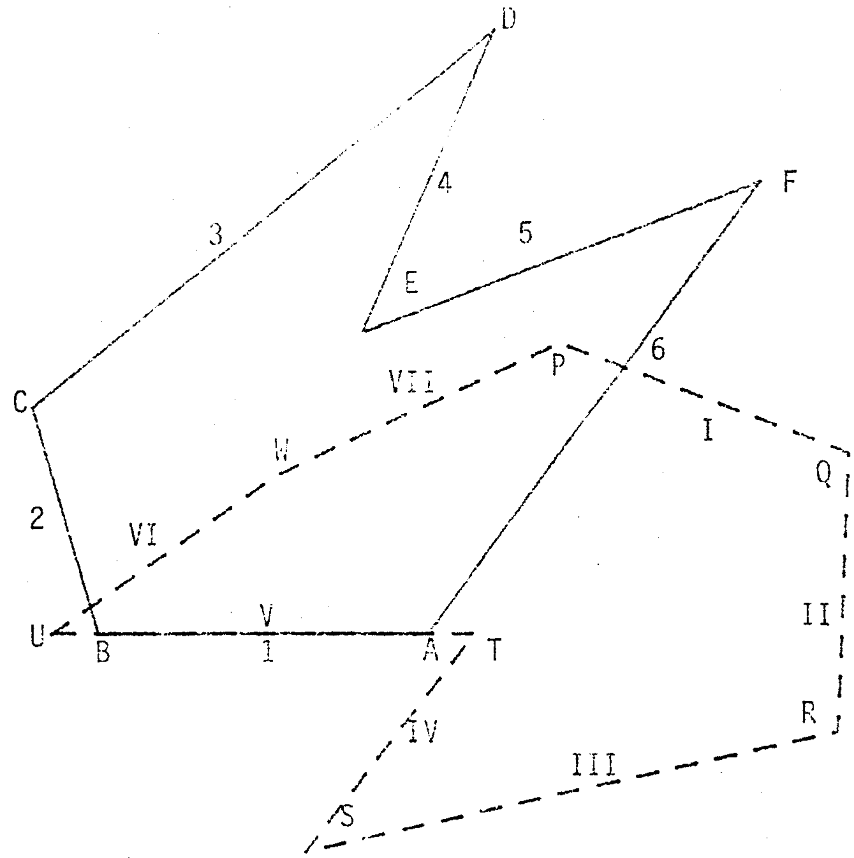

Fig. II.5. Computation of $D_{1,5}(\xi,\mu)$ between the hexagon ξ, ABCDEF, and the heptagon, μ, PQRSTUW. The edges of ξ and μ are given by the Arabic and Roman numerals respectively.

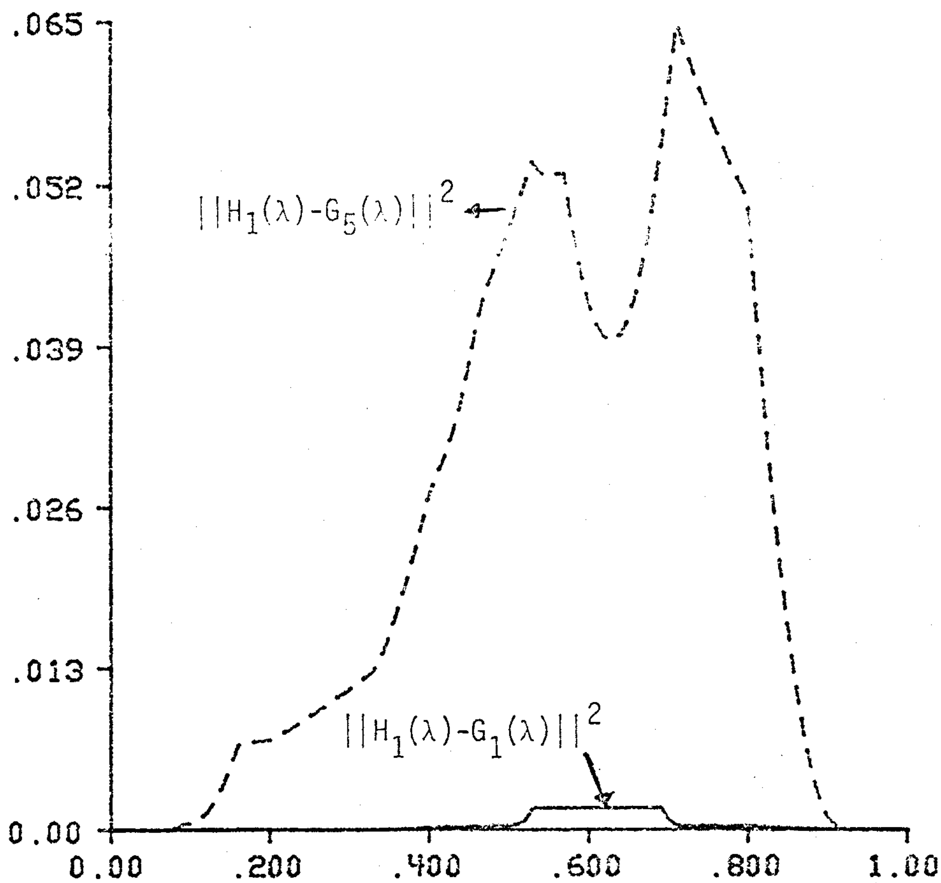

Fig. II.6. Plot demonstrating the capability of $||H_i(\lambda)-G_j(\lambda)||^2$ to capture the pointwise dissimilarity between the polygons ξ and μ of Example III.

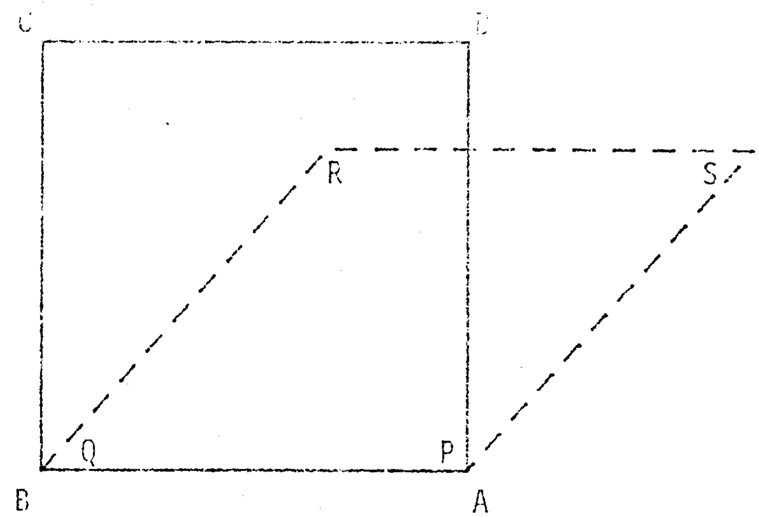

Fig. II.7. Computation of $D(\xi,\mu)$ between the square ξ, and the rhombus μ. The only configurations to be considered is the one when the edge AB of the square falls alongside the edge PQ of the rhombus.

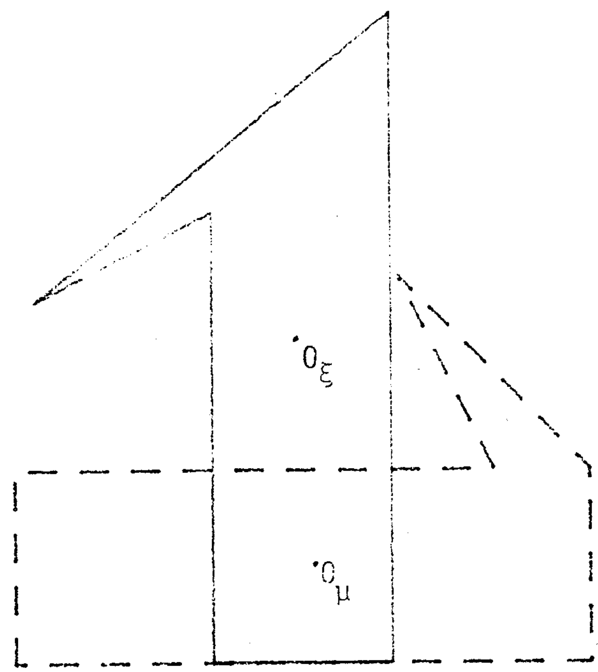

Fig. III.1. The distance between O_ξ and O_μ, the centers of gravity of ξ and μ, help to decide whether the integral square error in a particular configuration must be evaluated.

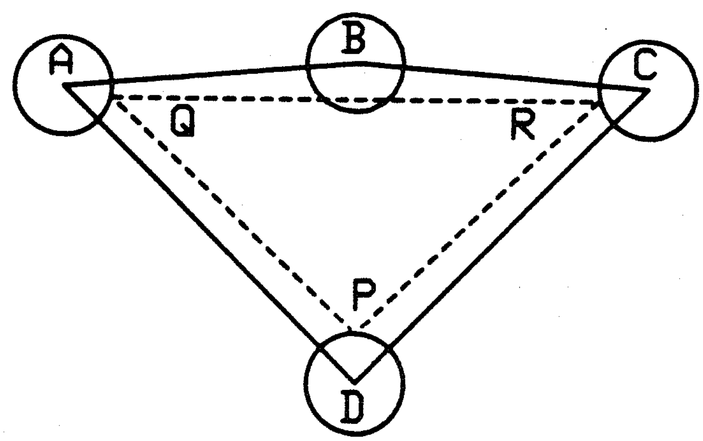

Fig. IV.1. The LMPP of the quadrilateral ABCD is the triangle PQR. The radii of the constraint disks are $\varepsilon(||AB|| + ||BC|| + ||CD|| + ||DA||)$.

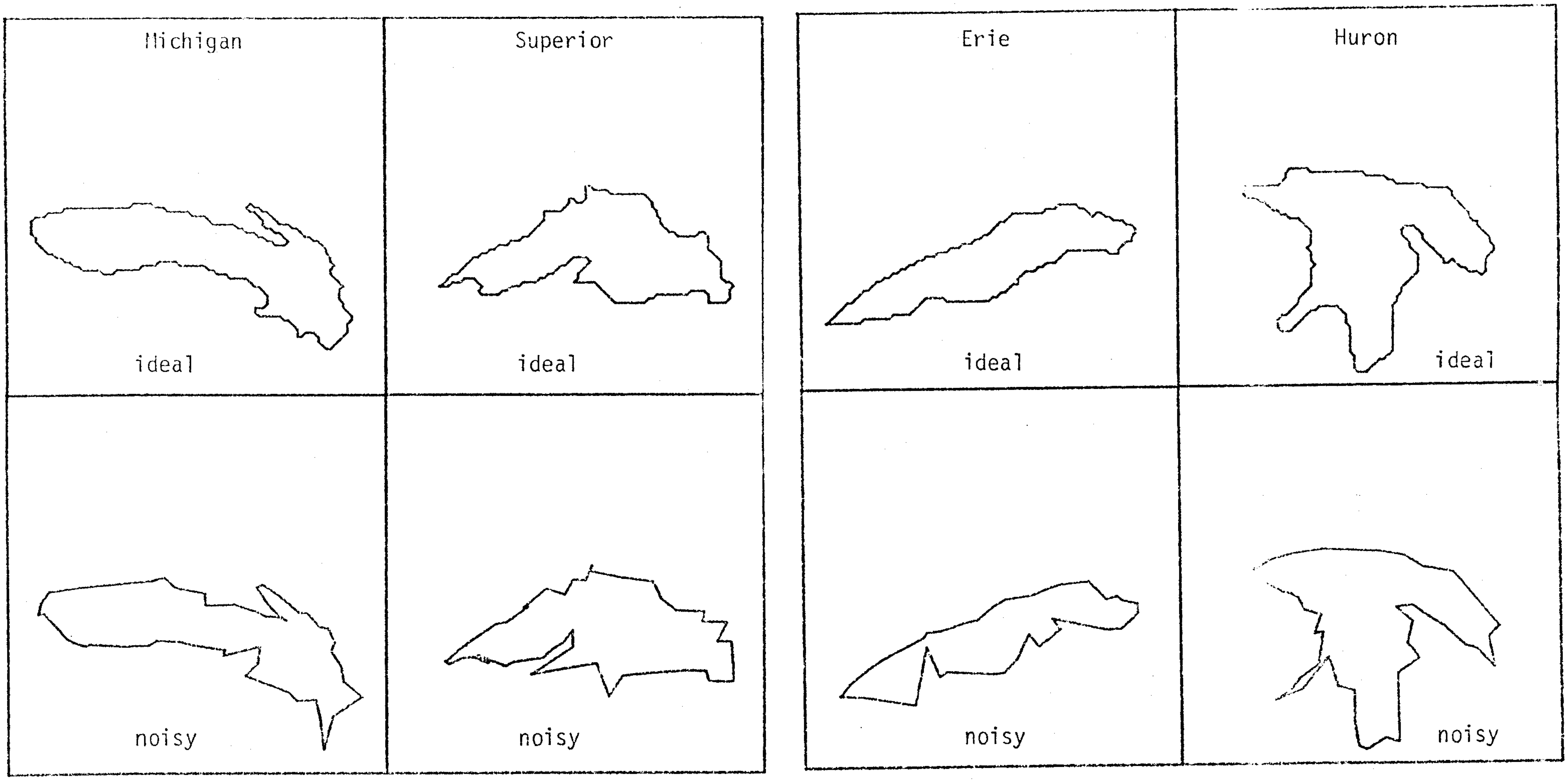

Figure V.1: Ideal and noisy boundaries of the 4 great lakes. Noise variance $\sigma^2 = 6.25$.

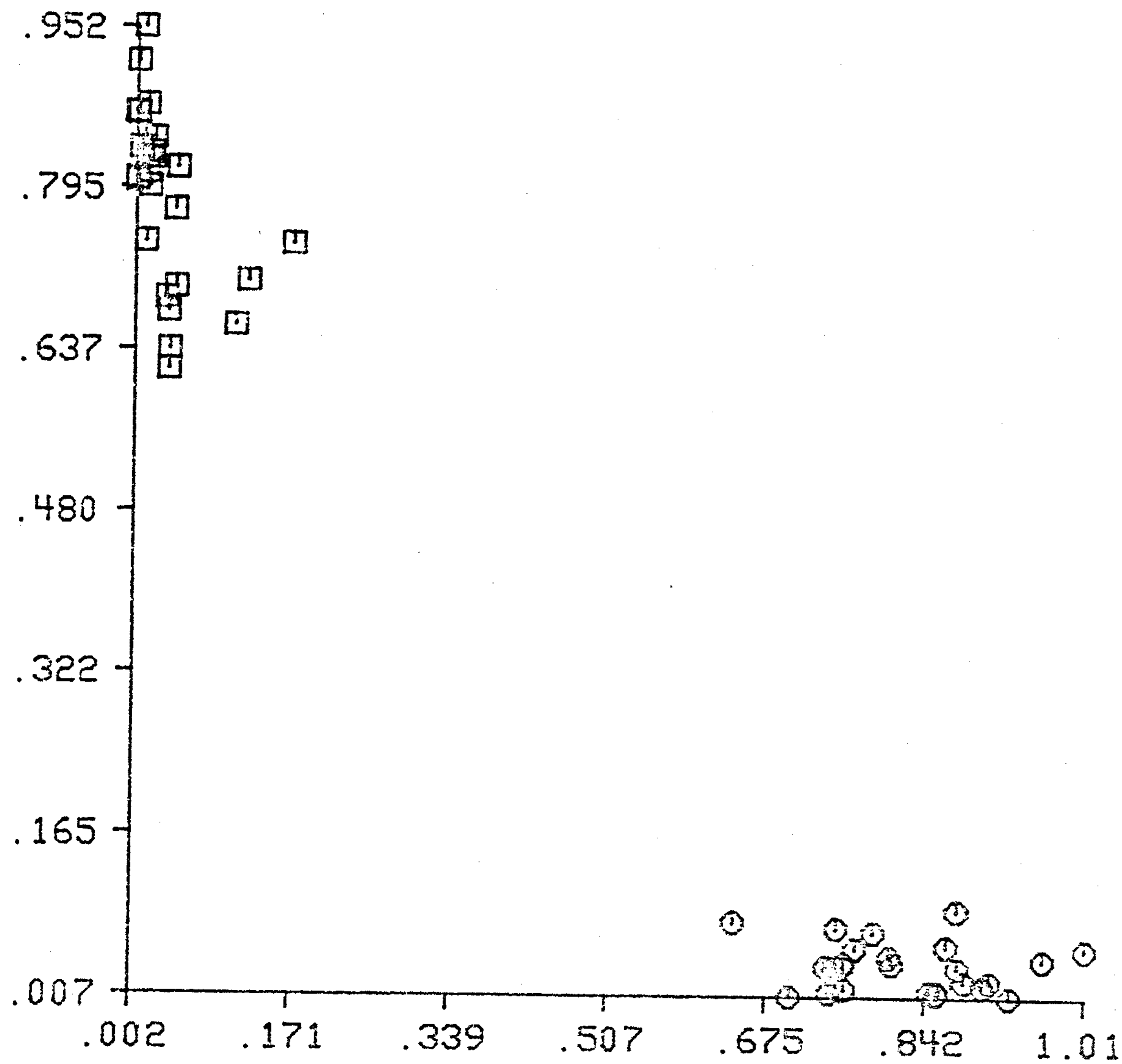

Fig. V.2. Graph of D(ideal Erie, μ) vs. D(ideal Huron, μ) for various noisy test samples, μ of Lake Erie, ⊡ , and Huron ⊙ .

CARLETON UNIVERSITY

School of Computer Science

BIBLIOGRAPHY OF SCS TECHNICAL REPORTS

SCS-TR-1 **THE DESIGN OF CP-6 PASCAL**
Jim des Rivieres and Wilf R. LaLonde, June 1982.

SCS-TR-2 **SINGLE PRODUCTION ELIMINATION IN LR(1) PARSERS: A SYNTHESIS**
Wilf R. LaLonde, June 1982.

SCS-TR-3 **A FLEXIBLE COMPILER STRUCTURE THAT ALLOWS DYNAMIC PHASE ORDERING**
Wilf R. LaLonde and Jim des Rivieres, June 1982.

SCS-TR-4 **A PRACTICAL LONGEST COMMON SUBSEQUENCE ALGORITHM FOR TEXT COLLATION**
Jim des Rivieres, June 1982.

SCS-TR-5 **A SCHOOL BUS ROUTING AND SCHEDULING PROBLEM**
Wolfgang Lindenberg, Frantisek Fiala, July 1982.

SCS-TR-6 **ROUTING WITHOUT ROUTING TABLES**
Nicola Santoro, Ramez Khatib, July 1982.

SCS-TR-7 **CONCURRENCY CONTROL IN LARGE COMPUTER NETWORKS**
Nicola Santoro, Hasan Ural, July 1982.

SCS-TR-8 Out of Print **ORDER STATISTICS ON DISTRIBUTED SETS**
Nicola Santoro, Jeffrey B. Sidney, July 1982.

SCS-TR-9 **OLIGARCHICAL CONTROL OF DISTRIBUTED PROCESSING SYSTEMS**
Moshe Krieger, Nicola Santoro, August 1982.

SCS-TR-10 **COMMUNICATION BOUNDS FOR SELECTION IN DISTRIBUTED SETS**
Nicola Santoro, Jeffrey B. Sidney, September 1982.

SCS-TR-11 **SIMPLE TECHNIQUE FOR CONVERTING FROM A PASCAL SHOP TO A C SHOP**
Wilf R. LaLonde, John R. Pugh, November 1982.

SCS-TR-12 **EFFICIENT ABSTRACT IMPLEMENTATIONS FOR RELATIONAL DATA STRUCTURES**
Nicola Santoro, December 1982.

SCS-TR-13 **ON THE MESSAGE COMPLEXITY OF DISTRIBUTED PROBLEMS**
Nicola Santoro, December 1982.

SCS-TR-14 **A COMMON BASIS FOR SIMILARITY MEASURES INVOLVING TWO STRINGS**
R. L. Kashyap and B. J. Oommen, January 1983.

SCS-TR-15 **SIMILARITY MEASURES FOR SETS OF STRINGS**
R. L. Kashyap and B. J. Oommen, January 1983.

SCS-TR-16 **THE NOISY SUBSTRING MATCHING PROBLEM**
R. L. Kashyap and B. J. Oommen, January 1983.

SCS-TR-17 **DISTRIBUTED ELECTION IN A CIRCLE WITHOUT A GLOBAL SENSE OF ORIENTATION**
E. Korach, D. Rotem, N. Santoro, January 1983.

SCS-TR-18 **A GEOMETRICAL APPROACH TO POLYGONAL DISSIMILARITY AND THE CLASSIFICATION OF CLOSED BOUNDARIES**
R. L. Kashyap and B. J. Oommen, January 1983.

SCS-TR-19 **SCALE PRESERVING SMOOTHING OF POLYGONS**
R. L. Kashyap and B. J. Oommen, January 1983.

SCS-TR-20 **NOT-QUITE-LINEAR RANDOM ACCESS MEMORIES**
Jim des Rivieres, Wilf LaLonde and Mike Dixon, August 1982, Revised March 1, 1983.

SCS-TR-21 **SHOUT ECHO SELECTION IN DISTRIBUTED FILES**
D. Rotem, N. Santoro, J. B. Sidney, March 1983.

SCS-TR-22 **DISTRIBUTED RANKING**
E. Korach, D. Rotem, N. Santoro, March 1983.

SCS-TR-23 **A REDUCTION TECHNIQUE FOR SELECTION IN DISTRIBUTED FILES : I**
N. Santoro, J. B. Sidney, April 1983.
Replaced by SCS-TR-69

SCS-TR-24 **LEARNING AUTOMATA POSSESSING ERGODICITY OF THE MEAN : THE TWO ACTION CASE**
M. A. L. Thathachar and B. J. Oommen, May 1983.

SCS-TR-25 **ACTORS - THE STAGE IS SET**
John R. Pugh, June 1983.

SCS-TR-26 **ON THE ESSENTIAL EQUIVALENCE OF TWO FAMILIES OF LEARNING AUTOMATA**
M. A. L. Thathachar and B. J. Oommen, May 1983.

SCS-TR-27 **GENERALIZED KRYLOV AUTOMATA AND THEIR APPLICABILITY TO LEARNING IN NONSTATIONARY ENVIROMENTS**
B. J. Oommen, June 1983.

SCS-TR-28 **ACTOR SYSTEMS: SELECTED FEATURES**
Wilf R. LaLonde, July 1983.

SCS-TR-29 **ANOTHER ADDENDUM TO KRONECKER'S THEORY OF PENCILS**
M. D. Atkinson, August 1983.

SCS-TR-30 **SOME TECHNIQUES FOR GROUP CHARACTER REDUCTION**
M. D. Atkinson and R. A. Hassan, August 1983.

SCS-TR-31 **AN OPTIMAL ALGORITHM FOR GEOMETRICAL CONGRUENCE**
M. D. Atkinson, August 1983.

SCS-TR-32 **MULTI-ACTION LEARNING AUTOMATA POSSESSING ERGODICITY OF THE MEAN**
B. J. Oommen and M. A. L. Thathachar, August 1983.

SCS-TR-33 **FIBONACCI GRAPHS, CYCLIC PERMUTATIONS AND EXTREMAL POINTS**
N. Santoro and J. Urrutia, December 1983

SCS-TR-34 **DISTRIBUTED SORTING**
D. Rotem, N. Santoro, and J. B. Sidney, December 1983

SCS-TR-35 **A REDUCTION TECHNIQUE FOR SELECTION IN DISTRIBUTED FILES: II**
N. Santoro, M. Scheutzow, and J. B. Sidney, December 1983

SCS-TR-36 **THE ASYMPTOTIC OPTIMALITY OF DISCRETIZED LINEAR REWARD-INACTION LEARNING AUTOMATA**
B. J. Oommen and Eldon Hansen, January 1984

SCS-TR-37 **GEOMETRIC CONTAINMENT IS NOT REDUCIBLE TO PARETO DOMINANCE**
N. Santoro, J. B. Sidney, S. J. Sidney, and J. Urrutia, January 1984

SCS-TR-38 **AN IMPROVED AGLORITHM FOR BOOLEAN MATRIX MULTIPLICATION**
N. Santoro and J. Urrutia, January 1984

SCS-TR-39 **CONTAINMENT OF ELEMENTARY GEOMETRIC OBJECTS**
J. Sack, N. Santoro and J. Urrutia, February 1984

SCS-TR-40 **SADE: A PROGRAMMING ENVIRONMENT FOR DESIGNING AND TESTING SYSTOLIC ALGORITHMS**
J. P. Corriveau and N. Santoro, February 1984

SCS-TR-41 **INTERSECTION GRAPHS, {B }-ORIENTABLE GRAPHS AND PROPER CIRCULAR ARC GRAPHS**
Jorge Urrutia, February 1984

SCS-TR-42 **MINIMUM DECOMPOSITIONS OF POLYGONAL OBJECTS**
J. Mark Keil and Jorg-R. Sack, March 1984

SCS-TR-43 **AN ALGORITHM FOR MERGING HEAPS**
Jorg-R. Sack and Thomas Strothotte, March 1984

SCS-TR-44 **A DIGITAL HASHING SCHEME FOR DYNAMIC MULTIATTRIBUTE FILES**
E. J. Otoo, March 1984

SCS-TR-45 **SYMMETRIC INDEX MAINTENANCE USING MULTIDIMENSIONAL LINEAR HASHING**
E. J. Otoo, March 1984

SCS-TR-46 **A MAPPING FUNCTION FOR THE DIRECTORY OF A MULTIDIMENSIONAL EXTENDIBLE HASHING**
E. J. Otoo, March 1984

SCS-TR-47 **TRANSLATING POLYGONS IN THE PLANE**
Jorg-R. Sack, March 1984

SCS-TR-48 **CONSTRAINED STRING EDITING**
J. Oommen, May 1984

SCS-TR-49 **O(N) ELECTION ALGORITHMS IN COMPLETE NETWORKS WITH GLOBAL SENSE OF ORIENTATION**
Jorg Sack, Nicola Santoro, Jorge Urrutia, May 1984

SCS-TR-50 **THE DESIGN OF A PROGRAM EDITOR BASED ON CONSTRAINTS**
Christopher A. Carter and Wilf R. LaLonde, May 1984

SCS-TR-51 **DISCRETIZED LINEAR INACTION-PENALTY LEARNING AUTOMATA**
B. J. Oommen and Eldon Hansen, May 1984

SCS-TR-52 **SENSE OF DIRECTION, TOPOLOGICAL AWARENESS AND COMMUNICATION COMPLEXITY**
Nicola Santoro, May 1984

SCS-TR-53 **OPTIMAL LIST ORGANIZING STRATEGY WHICH USES STOCHASTIC MOVE-TO-FRONT OPERATIONS**
B. J. Oommen, June 1984

SCS-TR-54 **RECTINLINEAR COMPUTATIONAL GEOMETRY**
J. Sack, June 1984

SCS-TR-55 **AN EFFICIENT, IMPLICIT DOUBLE-ENDED PRIORITY QUEUE**
M. D. Atkinson, Jorg Sack, Nicola Santoro, T. Strothotte, July 1984

SCS-TR-56 **DYNAMIC MULTIPAGING: A MULTIDIMENSIONAL STRUCTURE FOR FAST ASSOCIATIVE SEARCHING**
E. Otoo, T. H. Merrett

SCS-TR-57 **SPECIALIZATION, GENERALIZATION AND INHERITANCE**
Wilf R. LaLonde, John R. Pugh, August 1984

SCS-TR-58 COMPUTER ACCESS METHODS FOR EXTENDIBLE ARRAYS OF VARYING DIMENSIONS
E. Otoo, August 1984.

SCS-TR-59 AREA-EFFICIENT EMBEDDINGS OF TREES
J. P. Corriveau, Nicola Santoro, August 1984.

SCS-TR-60 UNIQUELY COLOURABLE m-DICHROMATIC ORIENTED GRAPHS
V. Neumann-Lara, N. Santoro, J. Urrutia, August 1984.

SCS-TR-61 ANALYSIS OF A DISTRIBUTED ALGORITHMS FOR EXTREMA FINDING IN A RING
D. Rotem, E. Korach and N. Santoro, August 1984.

SCS-TR-62 ON ZIGZAG PERMUTATIONS AND COMPARISONS OF ADJACENT ELEMENTS
M. D. Atkinson, October 1984

SCS-TR-63 SETS OF INTEGERS WITH DISTINCT DIFFERENCES
M. D. Atkinson, A. Hassenklover, October 1984.

SCS-TR-64 TEACHING FIFTH GENERATION COMPUTING: THE IMPORTANCE OF SMALL TALK
Wilf R. LaLonde, Dave A. Thomas, John R. Pugh, October 1984.

SCS-TR-65 AN EXTREMELY FAST MINIMUM SPANNING CIRCLE ALGORITHM
B. J. Oommen, October, 1984.

SCS-TR-66 ON THE FUTILITY OF ARBITRARILY INCREASING MEMORY CAPABILITIES OF STOCHASTIC LEARNING AUTOMATA
B. J. Oommen, October, 1984. Revised May 1985.

SCS-TR-67 HEAPS IN HEAPS
T. Strothotte, J.-R. Sack, November 1984. Revised April 1985.

SCS-TR-68 PARTIAL ORDERS AND COMPARISON PROBLEMS
M. D. Atkinson, November 1984.

SCS-TR-69 ON THE EXPECTED COMMUNICATION COMPLEXITY OF DISTRIBUTED SELECTION
N. Santoro, J. B. Sidney, S. J. Sidney, February 1985.

SCS-TR-70 FEATURES OF FIFTH GENERATION LANGUAGES: A PANORAMIC VIEW
Wilf R. LaLonde, John R. Pugh, March 1985.

SCS-TR-71 ACTRA: THE DESIGN OF AN INDUSTRIAL FIFTH GENERATION SMALLTALK SYSTEM
David A. Thomas, Wilf R. LaLonde, April 1985.

SCS-TR-72 MINMAXHEAPS, ORDERSTATISTICSTREES AND THEIR APPLICATION TO THE COURSEMARKS PROBLEM
M. D. Atkinson, J.-R. Sack, N. Santoro, Th. Strothotte, March 1985.

SCS-TR-73 DESIGNING COMMUNITIES OF DATA TYPES
Wilf R. LaLonde, May 1985.

SCS-TR-74 ABSORBING AND ERGODIC DISCRETIZED TWO ACTION LEARNING AUTOMATA
B. John Oommen, May 1985.

SCS-TR-75 OPTIMAL PARALLEL MERGING WITHOUT MEMORY CONFLICTS
Selim Akl and Nicola Santoro, May 1985.

SCS-TR-76 LIST ORGANIZING STRATEGIES USING STOCHASTIC MOVE-TO-FRONT AND STOCHASTIC MOVE-TO-REAR OPERATIONS
B. John Oommen, May 1985.