MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

Automatic Recognition of Regular Figures by Geometric AIC

Iman Triono† and Kenichi Kanatani†

Department of Computer Science, Gunma University

Abstract

We implement a graphical user interface that automatically transforms a figure input by a mouse into a regular figure which the system infers is the closest to the input. The difficulty lies in the fact that the classes into which the input is to be classified have inclusion relations among themselves. This prohibits us from using a simple distance criterion. In this paper, we show that this problem can be resolved by introducing the *geometric information criterion* (*AIC*) and give show implementation examples.

1. Introduction

A mouse is one of the most fundamental interfaces for generating figures interactively. One problem with it is generation of *regular* figures. Design of industrial objects and parts requires many kinds or regularity such as symmetry, orthogonality, and parallelism. However, the input figure need not satisfy the required regularity if the mouse is manipulated by a human operator. In many drawing tools, users are required to choose a specific mode of regularity (e.g., mode for rectangles) from a menu beforehand or enforce a specific regularity by inputting a command afterward. Is it not possible to automate this process? For example, if a user inputs an approximate rectangle or an approximate square, is it not possible for the computer to automatically infers the intended shape and correct the input figure into the inferred shape?

This appears simple at first sight. For example, we introduce some distance measure that describes dissimilarity between two figures, say the sum of the squared distances between the corresponding vertices. We prepare candidate classes of regular figures such as the class of rectangles and the class of squares. Given an input figure, we choose from each class the closest figure to the input in the distance measure we defined. Finally, we choose the one that has the smallest distance among them.

This paradigm has a fatal flaw. This is due to the fact that classes of regular figures have *inclusion relations*. For example, the class of squares is a subset of the class of rectangles. It follows that the distance from any figure to the closest square is always no more than the distance to the closest rectangle. This means that squares are never chosen. In general, classes that are included in other classes are never chosen whatever distance measure is used (Fig. 1).

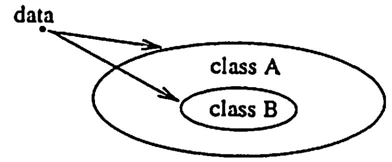


Figure 1: A class included in another is not chosen.

In *pattern recognition*, it is tacitly assumed that the classes into which an input is to be classified are *disjoint*. One solution to the problem of classes with inclusion is artificial separation of the classes. For example, we may introduce an artificial threshold ϵ and decide that a rectangle is a square if the ratio of the lengths of adjacent sides are between $1-\epsilon$ and $1+\epsilon$, thereby separating the class of squares from the class of rectangles. However, the inclusion relation is one of the most important bases of geometric reasoning; its artificial disruption might cause difficulties in automated reasoning. Moreover, how can we determine the threshold value ϵ? There exists no guiding principle for its determination.

In this paper, we present a framework that allows us to make judgement without destroying inclusion relations by introducing a criterion that favors a class included in another; judgment can be done by considering the balance of this criterion and the distance measure. A key idea is to use *statistical inference*.

In statistics, a well known criterion for selecting a reasonable model is what is called the *AIC* (*Akaike information criterion*) [1]. However, inferences in statistics are usually formulated as *estimating the parameters of the distribution from which the data are drawn*. It follows that it is difficult to apply the AIC to the problem we are now concerned. However, we can generalize the principle that underlies the AIC to geometric inference. The resulting criterion is called the *geometric AIC* [2, 3].

The guiding principle of the geometric AIC is choosing a model with high *predicting capability*. For example, let $\hat{S}$ be the closest square to an input figure Q. Let $\bar{S}$ be a square slightly different from the square $\hat{S}$, and let Q' be a quadrilateral slightly different from the square $\bar{S}$. Then, $\hat{S}$ is said to have predicting capability if it is also a good approximation to Q'. Similarly, let

†Address: Kiryu, Gunma 376, Japan.
E-mail: {iman|kanatani}@ail.cs.gunma-u.ac.jp.
This work was in part supported by the Ministry of Education, Science, Sports and Culture, Japan under the Grant in Aid for Scientific Research B (No. 07458067).

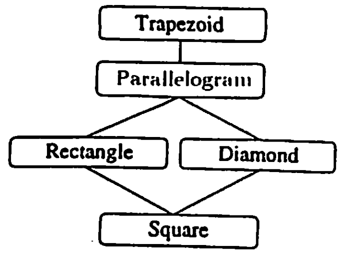


Figure 2: Class inclusion relations for quadrilaterals.

$\hat{R}$ is the closest rectangle to the input figure Q. Let $\bar{R}$ be a rectangle slightly different from the rectangle $\hat{R}$, and let Q' be a quadrilateral slightly different from the rectangle $\bar{R}$. Then, $\hat{R}$ has predicting capability if it is also a good approximation to Q'.

If the distance of Q to the square $\hat{S}$ and the distance of Q to the rectangle $\hat{R}$ is the same, the square $\hat{S}$ has higher predicting capability than the rectangle $\hat{R}$, because a rectangle has a larger degree of freedom than a square and hence the set $\mathcal{R}$ of rectangles that has a fixed distance to Q has a manifold of a higher dimensionality than the set $\mathcal{S}$ of squares that has the same distance to Q. It follows that perturbations from $\mathcal{S}$ constitute a smaller set than perturbations from $\mathcal{R}$, thereby making $\hat{S}$ a better approximation.

In the following, we formalize this intuitive argument in precise terms and present a scheme for classifying regular figures without introducing any threshold. Finally, we demonstrate the performance of the CAD system we have implemented on a workstation.

2. Classification of Quadrilaterals

Consider trapezoids, parallelograms, diamonds (rhombuses), rectangles, and squares as typical examples of regular figures (Fig. 2), although our analysis and procedure can apply to any other regular figures as well. We represent a point (x, y) in two dimensions by a three-dimensional vector $\boldsymbol{x} = (x, y, 1)^\top$ ($\top$ denotes transpose) and call a point with coordinates (x, y) simply "point $\boldsymbol{x}$". Consider a quadrilateral defined by connecting four points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_4$ in that order. The necessary and sufficient condition for it to be a regular figure listed in Fig. 2 is as follows. Here, $|\cdot, \cdot, |$ denotes scalar triple product, and $(\cdot, \cdot)$ denotes inner product; we define $\boldsymbol{k} = (0, 0, 1)^\top$.

Trapezoid. At least two sides are parallel:

$$|\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_3, \boldsymbol{k}| \cdot |\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_4, \boldsymbol{k}| = 0.$$

Parallelogram. The two pairs of sides are parallel:

$$|\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_3, \boldsymbol{k}| = 0, \quad |\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_4, \boldsymbol{k}| = 0.$$

Rectangle. The two pairs of sides are parallel, and adjacent sides are orthogonal:

$$|\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_3, \boldsymbol{k}| = 0, \quad |\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_4, \boldsymbol{k}| = 0,$$

$$(\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_2 - \boldsymbol{x}_1) = 0.$$

Diamonds (Rhombuses). The two pairs of sides are parallel, and the the two diagonals are orthogonal:

$$|\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_3, \boldsymbol{k}| = 0, \quad |\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_4, \boldsymbol{k}| = 0,$$

$$(\boldsymbol{x}_3 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_2) = 0.$$

Squares The two pairs of sides are parallel, adjacent sides are orthogonal, and the the two diagonals are orthogonal:

$$|\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_3, \boldsymbol{k}| = 0, \quad |\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_1 - \boldsymbol{x}_4, \boldsymbol{k}| = 0,$$

$$(\boldsymbol{x}_3 - \boldsymbol{x}_2, \boldsymbol{x}_2 - \boldsymbol{x}_1) = 0, \quad (\boldsymbol{x}_3 - \boldsymbol{x}_1, \boldsymbol{x}_4 - \boldsymbol{x}_2) = 0.$$

3. Optimal Correction

Consider the problem of correcting arbitrarily given N points $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_N$ into positions that satisfy given constraint. Suppose the constraint is given by L equations in the following form:

$$F^{(k)}(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) = 0, \qquad k = 1, ..., L. \tag{1}$$

Infinitely many ways exist to satisfy this constraint. Here, we correct the points by "minimal distances". As a measure of distance, we adopt the sum of squared distances over which the points are to be displaced. Let $\hat{\boldsymbol{x}}_1$, ..., $\hat{\boldsymbol{x}}_N$ be the corrected positions, and define

$$J = \sum_{\alpha=1}^{N} \|\boldsymbol{x}_\alpha - \hat{\boldsymbol{x}}_\alpha\|^2. \tag{2}$$

This quantity can be obtained by changing the sign of the logarithm of the likelihood (up to an additive constant) if the error of each point is subject to an independent Gaussian distribution with mean 0 and a constant variance. It follows that minimization of J can be interpreted to be *maximum likelihood estimation* in statistical terms.

Let $\Delta\boldsymbol{x}_1$, ..., $\Delta\boldsymbol{x}_N$ be the correction terms, and write

$$\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha - \Delta\boldsymbol{x}_\alpha. \tag{3}$$

Substituting this into eq. (1), expanding with respect to correction terms, assuming that they are small, and retaining only first order terms, we obtain

$$(\nabla_{\mathbf{x}_1} F^{(k)}, \Delta\boldsymbol{x}_1) + \cdots + (\nabla_{\mathbf{x}_N} F^{(k)}, \Delta\boldsymbol{x}_N) = F^{(k)}. \tag{4}$$

Here, $F^{(k)}$ is an abbreviation of $F^{(k)}(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$; $\nabla_{\mathbf{x}}(\cdot)$ denotes the gradient $(\partial(\cdot)/\partial x_1, ..., \partial(\cdot)/\partial x_N)^\top$ with respect to $\boldsymbol{x}$. Eq. (4) gives a set of L simultaneous linear equations in $\Delta\boldsymbol{x}_N$. We assume that the L equations (1) are mutually independent; we call L the *rank* of the constraint (1).

Introducing Lagrange multiplies, we obtain the solution that minimizes eq. (2) under the constraint (4) as follows:

$$\Delta\boldsymbol{x}_1 \oplus \cdots \oplus \Delta\boldsymbol{x}_N = \boldsymbol{A}^\top \boldsymbol{W} \boldsymbol{c}. \tag{5}$$

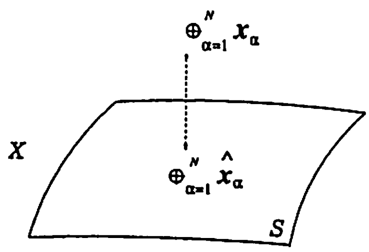


Figure 3: Orthogonal projection onto the model.

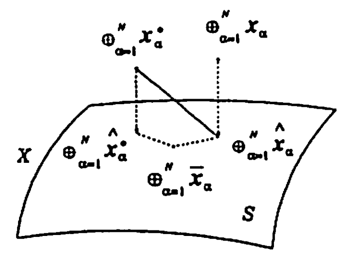


Figure 4: Derivation of geometric AIC.

The symbol $\oplus$ denotes direct product; $\boldsymbol{A}$, $\boldsymbol{W}$, and $\boldsymbol{c}$ are, respectively, an $L \times 3N$ matrix, an $L \times L$ matrix, and an L-dimensional vector defined as follows:

$$\boldsymbol{A} = \begin{pmatrix} \nabla_{\mathbf{x}_1} F^{(1)\top} & \cdots & \nabla_{\mathbf{x}_N} F^{(1)\top} \\ \vdots & \cdots & \vdots \\ \nabla_{\mathbf{x}_1} F^{(L)\top} & \cdots & \nabla_{\mathbf{x}_N} F^{(L)\top} \end{pmatrix}, \quad (6)$$

$$\boldsymbol{W} = (\boldsymbol{A}\boldsymbol{A}^\top)^{-1}, \quad \boldsymbol{c} = (F^{(1)}, \cdots, F^{(L)})^\top. \quad (7)$$

Since eq. (5) is obtained from eq. (4), which is a linear approximation of eq. (1), the solution does not necessarily satisfy eq. (1) exactly. So, we regard the corrected values $\hat{\boldsymbol{x}}_\alpha \leftarrow \boldsymbol{x}_\alpha - \Delta\boldsymbol{x}_\alpha$ as data $\boldsymbol{x}_\alpha$ and iterate the corrections until eq. (1) is sufficiently satisfied[1].

4. Geometric AIC

Since the third components of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ are all 1, they have $2N$ degrees of freedom in all. Hence, their direct product $\bigoplus_{\alpha=1}^N \boldsymbol{x}_\alpha$ can be identified with a point in a $2N$-dimensional space $\mathcal{X}$, in which eq. (1) defines a $(2N-L)$-dimensional manifold $\mathcal{S}$ with *codimension* L. We call $\mathcal{S}$ the *model* of eq. (1). Minimizing eq. (2) is equivalent to *orthogonally projecting* the data point onto the model $\mathcal{S}$ [2, 3] (Fig.3).

Let $\hat{J}$ be the *residual* of the minimization, i.e., the value of eq. (2) obtained by substituting the optimal solution for $\hat{\boldsymbol{x}}_\alpha$. If the data $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ are perturbed from the true positions $\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_N$ that satisfy eq. (1) by independent Gaussian noise of mean 0 and variance σ^2, it can be shown that $\hat{J}/\sigma^2$ is subject to a χ^2 distribution with L degrees in the first order [3]. This can be understood intuitively if we note that only the components of the noise along the L-dimensional "normal directions" to the model $\mathcal{S}$ contribute to the residual $\hat{J}$. Thus,

$$\hat{\sigma}^2 = \frac{\hat{J}}{L} \quad (8)$$

is an unbiased estimator of the variance σ^2.

If the noise in the true positions $\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_N$ are different, we would observe different data $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_N^*$. Let us call them the *future data*. Consider the residual with respect to the future data

$$J^* = \sum_{\alpha=1}^N \|\boldsymbol{x}_\alpha^* - \hat{\boldsymbol{x}}_\alpha\|^2. \quad (9)$$

[1]These are essentially the Newton iterations, so the convergence is quadratic in general. Hence, four or five iterations are sufficient for practical purposes.

We can infer that the corrected positions $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_N$ have *high predicting capability* if the expectation

$$I = E^*[E[J^*]] \quad (10)$$

is small, where $E^*[\,\cdot\,]$ and $E[\,\cdot\,]$ denote the expectation with respect to the future data $\{\boldsymbol{x}_\alpha^*\}$ and the current data $\{\boldsymbol{x}_\alpha\}$, respectively. We call eq. (10) the *expected residual*. Its unbiased estimator is given in the first order as follows [2, 3]:

$$AIC(\mathcal{S}) = \hat{J} + 2(2N-L)\sigma^2. \quad (11)$$

We call this the *geometric AIC* of model $\mathcal{S}$ [2, 3].

Derivation of eq. (11). Let $\bigoplus_{\alpha=1}^N \hat{\boldsymbol{x}}_\alpha^*$ be the orthogonal projection of the future data point $\bigoplus_{\alpha=1}^N \boldsymbol{x}_\alpha^*$ in the $2N$-dimensional space $\mathcal{X}$ onto the model $\mathcal{S}$ (Fig. 4). If we abbreviate the right-hand sides of eqs. (2) and (9) by $L^2_{\mathbf{X}\hat{\mathbf{X}}}$ and $L^2_{\hat{\mathbf{X}}\mathbf{X}^*}$, respectively, we have

$$I = E^*[L^2_{\mathbf{X}^*\hat{\mathbf{X}}^*}] + E^*[E[L^2_{\hat{\mathbf{X}}^*\hat{\mathbf{X}}}]]. \quad (12)$$

Since $\{\boldsymbol{x}_\alpha^*\}$ and $\{\boldsymbol{x}_\alpha\}$ have the same statistical characteristics, the first term on the right-hand side equals $E[\hat{J}]$. Let $\{\bar{\boldsymbol{x}}_\alpha\}$ be the true positions. Then, the second term on the right-hand side equals $E^*[L^2_{\hat{\mathbf{X}}^*\bar{\mathbf{X}}}] + E[L^2_{\bar{\mathbf{X}}\hat{\mathbf{X}}}]$. Since the model $\mathcal{S}$ is $(2N-L)$-dimensional and since $\bigoplus_{\alpha=1}^N \hat{\boldsymbol{x}}_\alpha^*$ and $\bigoplus_{\alpha=1}^N \hat{\boldsymbol{x}}_\alpha$ are mutually independent in $\mathcal{S}$, quantities $L^2_{\hat{\mathbf{X}}^*\bar{\mathbf{X}}}/\sigma^2$ and $L^2_{\bar{\mathbf{X}}\hat{\mathbf{X}}}/\sigma^2$ are subject to independent χ^2 distributions with $2N-L$ degrees of freedom in the first order. Thus, each has expectation $2N-L$. Hence, we obtain eq. (11). □

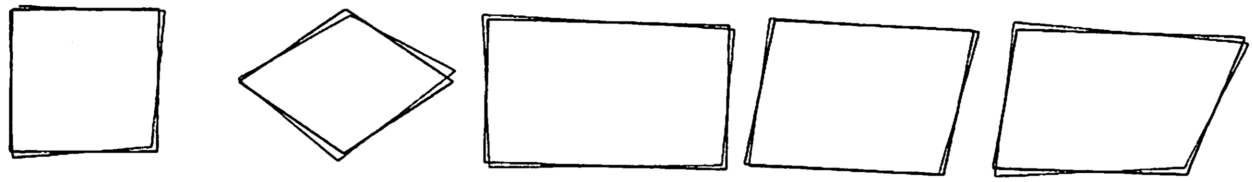
Consider another model $\mathcal{S}'$ with rank L', and let $\hat{J}'$ be the residual for this model. Its geometric AIC is given by

$$AIC(\mathcal{S}') = \hat{J}' + 2(2N-L')\sigma^2. \quad (13)$$

If $AIC(\mathcal{S}') < AIC(\mathcal{S})$, model $\mathcal{S}'$ is expected to have higher predicting capability than model $\mathcal{S}$. The condition for this can be written as follows:

$$\hat{J}' - \hat{J} < 2(L'-L)\sigma^2. \quad (14)$$

In order to apply eq. (14) for model selection, the variance σ^2 of the noise must be known. This is intuitively evident. In fact, suppose we want to distinguish squares from rectangles. A small difference has a significant meaning if the noise is known to be small,



(a) (b) (c) (d) (e)

Figure 5: Input figures and inferred shapes: (a) trapezoid; (b) parallelogram; (c) rectangle; (d) diamond; (e) square.

while such a distinction is meaningless if the noise is known to be large.

In general, we need other sources of information (e.g, an empirical distribution of the noise) to estimate the noise magnitude. If models have inclusion relations, however, we can take advantage of that fact to extract information about the noise. Suppose the L' equations that define the constraint on model $\mathcal{S}'$ can be obtained by adding some equations to the L equations that define the constraint on model $\mathcal{S}$. Then, we say model $\mathcal{S}'$ is *stronger* than model $\mathcal{S}$, or model $\mathcal{S}$ is *weaker* than model $\mathcal{S}'$, and write

$$\mathcal{S}' \succ \mathcal{S}. \tag{15}$$

Geometrically, this means that $\mathcal{S}'$ is a *submanifold* of $\mathcal{S}$ in $\mathcal{X}$. Suppose the weaker model $\mathcal{S}$ holds. Then, the variance σ^2 of the noise can be estimated by eq. (8) whether or not the stronger model $\mathcal{S}'$ is satisfied. Substituting the estimate into eq. (14), we obtain the following criterion:

$$\frac{J'}{J} < \frac{2L' - L}{L}. \tag{16}$$

5. Classification Procedure

Let $\hat{J}_{\rm trap}$ be the residual for optimally correcting four input points x_1, x_2, x_3, x_4 into vertices of a trapezoid. Similarly, let $\hat{J}_{\rm para}$, $\hat{J}_{\rm rect}$, $\hat{J}_{\rm diam}$, and $\hat{J}_{\rm squa}$ be the residuals for correcting the input into a parallelogram, a rectangle, a diamond, and a square, respectively. The procedure for discrimination is as follows:

1. (a) If $\hat{J}_{\rm para} \leq 3\hat{J}_{\rm trap}$, judge the input figure to be a parallelogram and go to Step 2.
   (b) Output an optimally corrected trapezoid, and stop.
2. (a) If $\hat{J}_{\rm rectangle} \leq 2\hat{J}_{\rm para}$, judge the input figure to be a rectangle and go to Step 3.
   (b) If $\hat{J}_{\rm diam} \leq 2\hat{J}_{\rm para}$, judge the input figure to be a diamond and go to Step 4.
   (c) Output an optimally corrected parallelogram, and stop.
3. (a) If $\hat{J}_{\rm squa} \leq (5/3)\hat{J}_{\rm rect}$, judge the input figure to be a square, output an optimally corrected square, and stop.
   (b) Output an optimally corrected rectangle, and stop.
4. (a) If $\hat{J}_{\rm squa} \leq (5/3)\hat{J}_{\rm diam}$, judge the input figure to be a square, output an optimally corrected square, and stop.
   (b) Output an optimally corrected diamond, and stop.

Note that *this procedure does not involve any arbitrarily adjustable threshold* ϵ.

Figure 5 shows implementation examples: input figures are inferred to be (a) a trapezoid, (b) a parallelogram, (c) a rectangle, (d) a diamond, and (e) a square. The input figure is drawn in thin lines; the corrected figure is drawn in thick lines.

6. Concluding Remarks

In this paper, we have presented a graphic interface for inferring regularity in a figure input by a mouse without using any adjustable thresholds and automatically imposing the inferred regularity. We first argued that, unlike the usual pattern recognition problem, we cannot select the closest class measured by some distance criterion if the classes have *inclusion relations*. Then, we showed that this difficulty can be resolved by introducing the *geometric AIC*. We described the basic principle underlying it and demonstrated the system performance.

Although recognition, extraction, and classification of features that have inclusion relations have been studied in relation to a variety of applications [5, 6], the intrinsic difficulty caused by the inclusion relations does not seem to be fully understood yet [4]. The approach presented here is expected to play an important role in dealing such problems.

References

[1] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Automation Control*, 19-6 (1974), 176–723.

[2] K. Kanatani, Automatic singularity test for motion analysis by an information criterion, *Proc. 4th European Conf. Comput. Vision*, April 1996, Cambridge, U.K., Vol. 1, pp. 697–708.

[3] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, 1996.

[4] K. Kanatani, Comments on "Symmetry as a continuous feature", *IEEE Trans. Patt. Anal. Mach. Intell.*, **18** (1996), to appear.

[5] P. L. Rosin and G. A. West, Nonparametric segmentation of curves into various representations, *IEEE Trans. Patt. Anal. Mach. Intell.*, 17-12 (1995), 1140–1153.

[6] H. Zabrodsky, S. Peleg and D. Avnir, Symmetry as a continuous feature, *IEEE Trans. Patt. Anal. Mach. Intell.*, 17-12 (1995), 1154–1166.