Comparison of Combined Shape Descriptors for Irregular Objects

Jukka Iivarinen, Markus Peura, Jaakko Särelä, and Ari Visa[†]

Helsinki University of Technology
Laboratory of Computer and Information Science
P.O. Box 2200, FIN-02015 HUT, Finland
{Jukka.Iivarinen, Markus.Peura, Jaakko.Sarela}@hut.fi

[†]Lappeenranta University of Technology
Department of Information Technology
P.O. Box 20, FIN-53851 Lappeenranta, Finland
Ari.Visa@lut.fi

Abstract

This paper focuses on recognition powers and computational efforts of three different shape coding techniques, namely the chain code histogram (CCH), the pairwise geometric histogram (PGH), and the combination of simple shape descriptors, for characterization of irregular objects. In recognizing irregular objects the essential task is to design efficient measures based on relatively small prior knowledge on geometrical constraints of possible target objects. Three rather different approaches are evaluated and discussed by the means of the self-organizing map (SOM). A database retrieval problem is also assumed to further test their discriminatory powers. As a case study, natural irregular objects have been used. Grouping of these objects based on their visual similarity is the main topic in this paper. The combination of simple shape descriptors is shown to have good recognition capabilities and low computation costs.

1 Introduction

Recognition of objects is one of the basic tasks in computer vision applications. Recognition is usually based on gray levels or colors, and shape characteristics of target objects. The goal of object recognition is to find a description which contains sufficient information to distinguish between different target objects. The design of potential target objects is often known in advance, so the geometric used in creating the object (i.e., in manufacturing or in some natural formation process) can be directly applied to the recognition task. Consequently, the major difficulties might be merely in arranging the physical image acquisition. Plenty of methods exist for recognition of objects [14, 12, 17]. However, some of the methods, for example syntactical methods, are mainly suitable for regular or man-made object recognition. In this paper the main concern is on irregular objects (for example surface defects) which are hard to recognize even for a human observer.

There are several methods for the shape analysis of objects [14, 12, 17]. These methods can be divided into two categories, the area-based methods and the contour-based methods. The latter methods are of interest in this paper since we are mainly concerned on the shape of a contour. The contour-based methods include the following techniques. Simple descriptors, for example perimeter length, curvature, and bending energy, have been applied widely [12, 17]. Moment-based techniques have been used in object recognition since 1962 [8]. Moments derived from the contour of an object were used by Dudani et al. [3] and Gupta and Srinath [7]. Zahn and Roskies [18] used the Fourier coefficients of a contour as shape descriptors. The chord distribution of a contour was proposed by Smith and Jain [16]. A scale-space technique to form a description for plane curves was proposed by Mokhtarian and Mackworth [13]. Evans et al. [4] proposed pairwise geometric histograms as shape descriptors. The chain code histogram was proposed by Iivarinen and Visa [10].

In Section 2 three different shape coding techniques are shortly introduced. The chain code histogram (CCH) is a statistical measure for the directionality of the contour of an object [10]. It is calculated from the chain code presentation of the contour. The pairwise geometric histogram (PGH) was introduced by Evans et al. [4]. Lately Ashbrook et al. developed the algorithm further [2]. The PGH is a shape descriptor which is applied to polygonal objects. The last technique uses simple shape descriptors [12, 17]. The combination of five such descriptors seem to provide good object recognition capabilities [15]. In Section 3 the discriminatory powers of these three shape coding techniques are demonstrated on real irregular objects. Further discussions on the properties of these techniques are provided in Section 4. Conclusions are drawn in Section 5.

2 Shape Coding Techniques

Three shape coding techniques are shortly introduced. All of them are computed from the contour of an object. The object is expected to be the ordered (x, y) presentation of the contour pixels.

2.1 The Chain Code Histogram

The chain code histogram (CCH) is meant to group together objects that look similar to a human observer [10]. It is not meant for exact detection and classification tasks. The CCH is calculated from the chain code presentation of a contour.

The Freeman chain code [6] is a compact way to represent a contour of an object. The chain code is an ordered sequence of n links $\{c_i,\ i = 1, 2, ..., n\}$, where c_i is a vector connecting neighboring contour pixels. The directions of c_i are coded with integer values $k = 0, 1, ..., K-1$ in a counterclockwise sense starting from the direction of the positive x-axis. The number of directions K takes integer values $2^{(M+1)}$ where M is a positive integer. The chain codes where $K > 8$ are called generalized chain codes [5].

The calculation of the chain code histogram is fast and simple. The CCH is a discrete function

$$p(k) = \frac{n_k}{n}, \quad k = 0, 1, ..., K-1, \tag{1}$$

where n_k is the number of chain code values k in a chain code, and n is the number of links in a chain code. A simple example is depicted in Figure 1. In Figure 1(a) are the directions of the eight connected chain code. In Figure 1(b) is a sample object, a square. The starting point for the chain coding is marked with a black circle, and the chain coding direction is clockwise. In Figures 1(c)-(d) are the chain code and the CCH of the contour of the square.

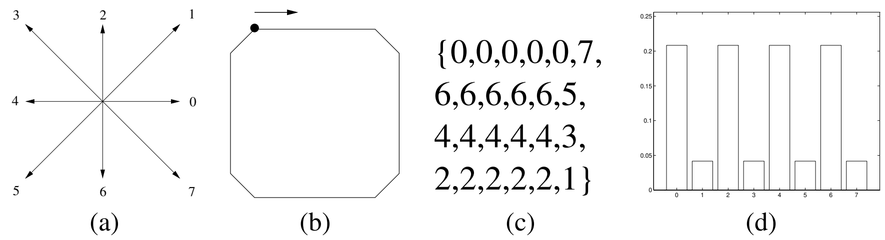


Figure 1: (a) The directions of the eight connected chain code ($K = 8$). (b) A sample object, a square, (c) chain code presentation of the square, and (d) the chain code histogram of the square.

The CCH is a translation and scale invariant shape descriptor. It can be made invariant to rotations of $90°$ because the $90°$ rotation causes only a circular shift in the CCH. To achieve better rotation invariance the normalized chain code histogram (NCCH) should be used [10]. It takes into account the lengths of different directions.

2.2 The Pairwise Geometric Histogram

The pairwise geometric histogram (PGH) is a powerful shape descriptor that is applied to polygonal shapes [4, 2]. It can be applied also to an irregular shape if the shape is first approximated with a polygon. For a recent review of polygon approximation algorithms, see for example [19].

Consider a polygon defined by its edgepoints $(\tilde{x}(t), \tilde{y}(t)) \in \mathcal{R}^2$. Now successive edgepoints define the line segments the polygon consists of. The PGH is calculated using the following strategy: Let each line segment be a reference line on its turn. Then the relative angle $\theta \in [0, \pi[$ and the perpendicular minimum and maximum distances (d_{min} and d_{max}) are calculated between the reference line and all the other lines, as shown in Figure 2(a). The histogram values are increased by one on the indexes corresponding to the angle θ and the line segment from the d_{min} to d_{max} (Figure 2(b)).

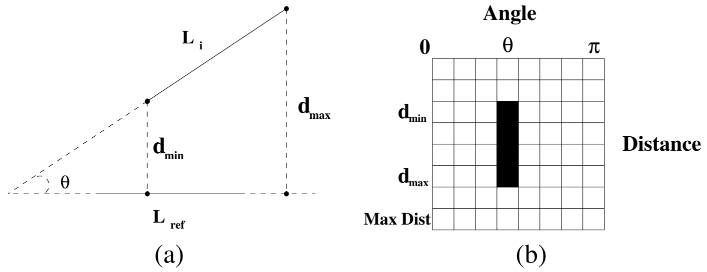


Figure 2: (a) Relative angle and perpendicular distances between two lines, and (b) the pairwise geometric histogram.

A new scheme to reduce the size of the PGH is proposed. Conditional expectations of each row and column are calculated and collected into a feature vector. Let $p(i,j)$ be the PGH value in the position (i,j). Then the feature vector $\mathbf{f}_{PGH}$ is given as

$$\mathbf{f}_{PGH} = [E_r(1)\ E_r(2)\ ...\ E_r(N)\ E_c(1)\ E_c(2)\ ...\ E_c(M)]^T, \tag{2}$$

where N is the number of rows of the PGH, M is the number of columns of the PGH, $E_r(i)$ is the conditional expectation of the ith row,

$$E_r(i) = \frac{\sum_j jp(i,j)}{\sum_j p(i,j)}, \tag{3}$$

and $E_c(j)$ is the conditional expectation of the jth column,

$$E_c(j) = \frac{\sum_i ip(i,j)}{\sum_i p(i,j)}. \tag{4}$$

Thus the number of features is reduced from NM to $N + M$. This presents significant savings in calculation time and memory requirements, especially if N and M are large.

2.3 Combination of Simple Shape Descriptors

In this section, five simple shape descriptors are introduced (Figure 3). Variations of most of them have been widely used in object recognition [12, 17]. Each descriptor alone is insufficient for a complex recognition task, but the combination of them is shown to have good recognition capabilities (Section 3). All these shape descriptors require only $\mathcal{O}(N)$ calculation steps expect convexity which requires $\mathcal{O}(N^2)$ calculation steps. For a more profound discussion and definitions of these descriptors, see [15].

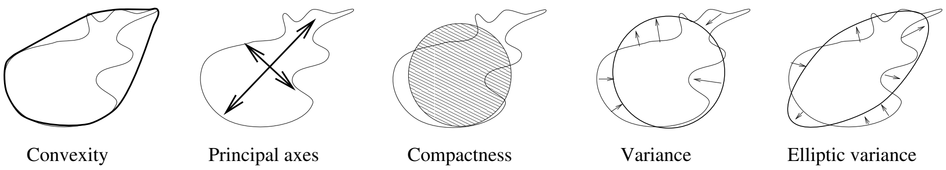


Figure 3: Five simple shape descriptors.

Convexity can be defined as the ratio of perimeters of the convex hull of the contour and the original contour. The convex hull is the minimal convex covering of an object. The algorithm for constructing a convex hull involves traversing the contour and minimizing turn angle in each step. Practically, dot product can be maximized instead.

Principal axes of an object can be uniquely defined as segments of lines crossing each other orthogonally in the centroid of the object and representing the directions with zero cross-correlation. This way, a contour is seen as an realization of a statistical distribution. The ratio of principal axis can be calculated from the covariance matrix of a contour. It is not necessary to calculate actual eigenvectors nor eigenvalues.

Compactness is often defined as the ratio of squared perimeter and the area of an object. It reaches the minimum in a circular object and approaches infinity in thin, complex objects. The measure applied in this paper is the ratio of the perimeter of a circle with equal area as the original object and the original perimeter.

Sometimes a shape should be compared against a template. A circle is an obvious and general template choice. The circular variance is the proportional mean-squared error with respect to solid circle. It gives zero for a perfect circle and increases along shape complexity and elongation.

Elliptic variance is defined similarly to the circular variance. An ellipse is fitted to the shape (instead of a circle) and the mean-squared error is measured.

3 Experiments

The discriminatory powers of the three shape coding techniques are demonstrated on a data set that consists of 79 irregular objects (Figure 4). These objects are surface defects which are obtained from base paper samples via an unsupervised segmentation procedure [9]. The distribution of different shapes does not necessarily possess any distinct clusters. That is, one cannot tell when a class changes to another and any classification becomes impossible. More likely, different shape classes vary smoothly and are more suitably interpreted by continuous descriptors.

Each object was given as an ordered (x, y) presentation of the contour pixels. In case of the chain code histogram (CCH), each contour was first chaincoded with an 8-connected chain code and then the CCH was calculated. When forming the pairwise geometric histogram (PGH) the contours of all objects were first approximated with polygons. The curvature guided polygonal approximation algorithm proposed in [1] was used. The number of angles and distances was quantized to 8 levels. The 8×8 PGH was thus obtained. To further reduce the size of the PGH, conditional expectations of each row and column were calculated and collected to a 16-dimensional feature vector. In case of the five simple shape descriptors, each descriptor was calculated, normalized to $[0, 1]$, and put into a feature vector.

The self-organizing map (SOM)[11] is used to visualize the ordering of objects. In Figures 5(a)-(c) are the SOMs which are trained with the three shape coding techniques. The previously mentioned data set of 79 irregular objects is used in training. If many objects are mapped to one map unit, only one object is depicted. All shape coding techniques seem to produce a logical ordering of objects. However, the ordering of objects with the chain code histogram is not as good as with other methods since the CCH cannot distinct between smooth and unsmooth objects. The pairwise geometric histogram and the simple shape descriptors seem to produce a very similar ordering. Smooth transitions from smooth to unsmooth objects and from circle-like to rod-like objects are nicely illustrated.

To further test the discriminatory powers of the three shape coding techniques a database retrieval problem is assumed. Three different shaped objects were selected as prototypes and the database (which consisted of 76 objects) was searched for three most similar objects. The prototypes and the retrieved objects are depicted in Figures 6(a)-(c). The nearest retrieved object is at the top and the third nearest at the bottom of each column. The prototype object A is a thin rod-like object, the prototype object B is a smooth oval object, and the prototype object C is an unsmooth object. For the prototype object A the results are quite similar. Unlike the other descriptors the chain code histogram is

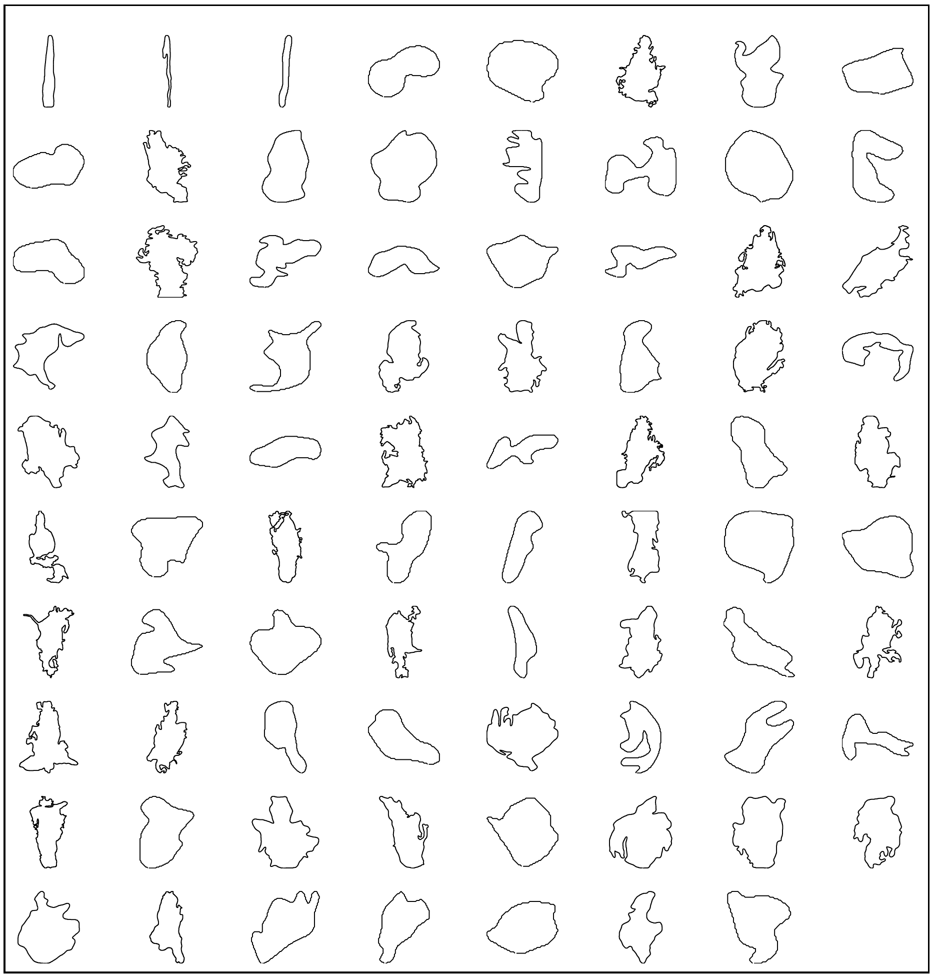


Figure 4: A set of 79 irregular objects used in experiments.

not a rotation invariant descriptor. This results some differences in the results. For the prototype objects B and C the results are similar with the pairwise geometric histogram and with the simple shape descriptors. The chain code histogram cannot distinct between smooth and unsmooth objects.

4 Discussion

The choice of the proper shape recognition method is always a compromise between recognition power and computational complexity. Speed requirements of real-time applications often limit the number of possible techniques. In previous chapter the recognition powers of three shape coding techniques were demonstrated. The pairwise geometric histogram and the set of five simple shape descriptors performed well, but the chain code histogram had some limitations. The CCH cannot preserve information on the exact shape

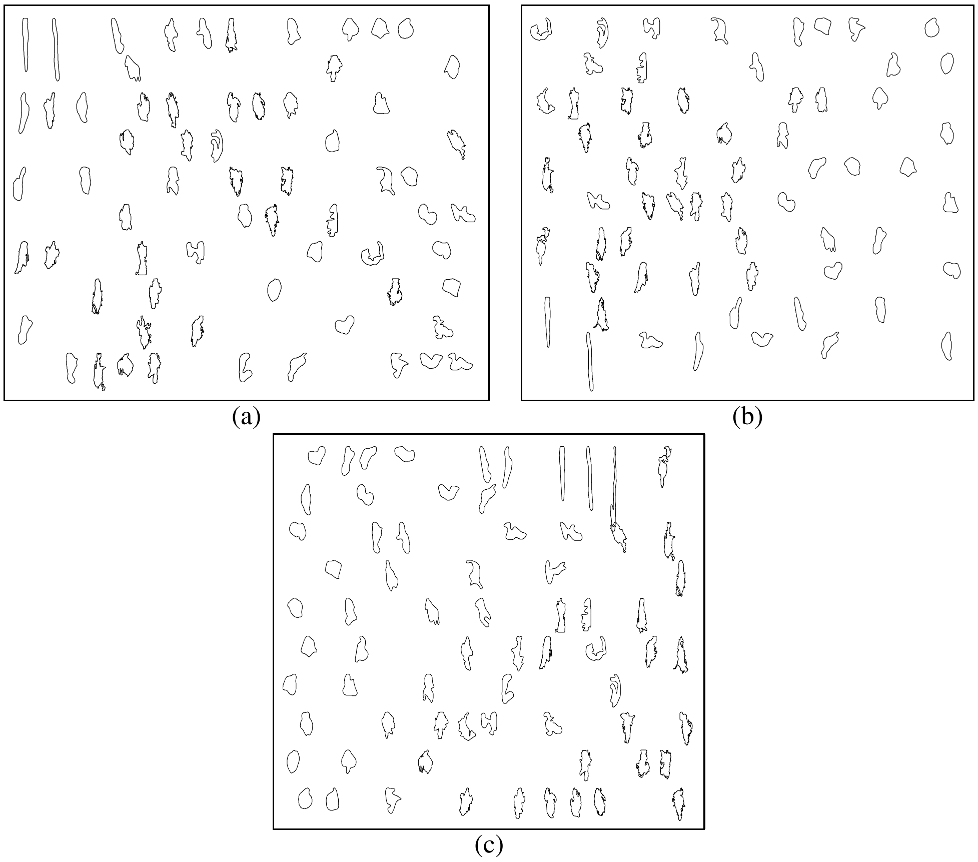


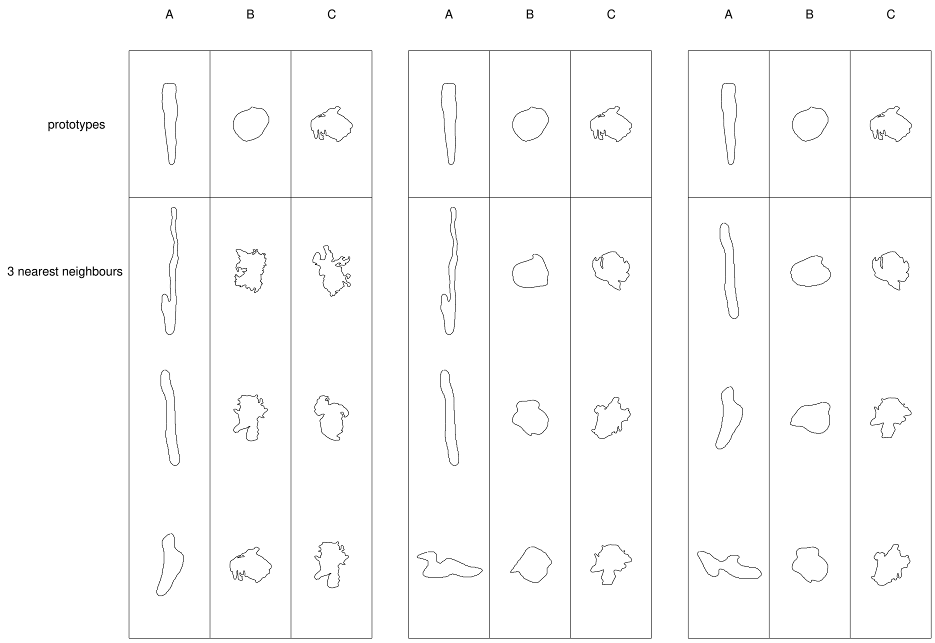
Figure 5: Sample contours on a self-organizing map. (a) The chain code histogram, (b) the pairwise geometric histogram, and (c) the set of five simple shape descriptors is used.

of a contour, because it only shows the probabilities (or frequencies) for the different directions present in a contour. Thus, there may be many objects with the same CCH. However, the chain code histogram is fast to calculate and it needs only a small amount of memory which can be crucial in many real-time applications. The CCH is evidently good enough for many applications, especially for those having relatively distinct classes.

The pairwise geometric histogram is computationally heavy and it requires more memory than the CCH. If objects are non-polygonal, the polygonal approximation must be made which requires more time. The PGH is suitable for problems that contain (nearly) polygonal objects. In addition, these objects can also be partially occluded.

The set of five simple shape descriptors fall in between the CCH and the PGH in both time and memory requirements. They seem to provide consistent recognition which agrees well with that of a human observer. Some of these descriptors are quite correlated [15], so only a subset of them may be sufficient in most applications.

The self-organizing map (SOM) is an ideal tool in exploring the natural structure of high-dimensional data. It makes a topology-preserving mapping from data space to 2D space. Practically all statistically significant information is preserved in the mapping.



(a) (b) (c)

Figure 6: Three most similar objects for three prototype objects retrieved from a database. (a) The chain code histogram, (b) the pairwise geometric histogram and (c) the set of five simple shape descriptors is used.

The SOM can discover features that are not explicitly present in single measurements. For example, the contours in Figure 5(c), which were ordered by the SOM according to five descriptors, seem to have roughly three dimensions, elongation, smoothness, and bending, none of which were obtainable using only one of the original descriptors.

5 Conclusions

Three different approaches to shape recognition were evaluated. The recognition powers were demonstrated with real irregular objects and computational aspects were discussed. The chain code histogram is a simple descriptor which is fast to calculate. However, it has some limitations which affect its generality. The pairwise geometric histogram is a very powerful descriptor which can be used especially with polygonal objects that can even be overlapping. It is, however, very computing intensive and requires much memory. The set of five simple shape descriptors provides a good compromise between recognition power and computational complexity. It performs equally well with the PGH but requires less computation time and memory.

Acknowledgments

The authors wish to thank the Technology Development Centre of Finland for financial support (TEKES's grant 4172/95).

References

[1] Nirwan Ansari and Edward J. Delp. On detecting dominant points. *Pattern Recognition*, 24:441–451, 1991.

[2] A. P. Ashbrook, N. A. Thacker, and P. I. Rockett. Pairwise geometric histograms: A scaleable solution for the recognition of 2d rigid shape. In *Proceedings of the 9th Scandinavian Conference on Image Analysis*, volume 1, pages 271–278, Uppsala, Sweden, June 1995.

[3] S. A. Dudani, K. J. Breeding, and R. B. McGhee. Aircraft identification by moment invariants. *IEEE Transactions on Computers*, C-26:39–46, 1977.

[4] Alun C. Evans, Neil A. Thacker, and John E. W. Mayhew. Pairwise representations of shape. In *Proceedings of the 11th IAPR International Conference on Pattern Recognition*, volume 1, pages 133–136, The Hague, The Netherlands, August 30-September 3 1992.

[5] H. Freeman and A. Saghri. Generalized chain codes for planar curves. In *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pages 701–703, Kyoto, Japan, November 7-10 1978.

[6] Herbert Freeman. Computer processing of line-drawing images. *Computing Surveys*, 6(1):57–97, March 1974.

[7] L. Gupta and M. D. Srinath. Contour sequence moments for the classification of closed planar shapes. *Pattern Recognition*, 20(3):267–272, 1987.

[8] Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, IT-8:179–187, 1962.

[9] Jukka Iivarinen, Juhani Rauhamaa, and Ari Visa. Unsupervised segmentation of surface defects. In *13th International Conference on Pattern Recognition*, volume IV, pages 356–360, Wien, Austria, August 25–30 1996.

[10] Jukka Iivarinen and Ari Visa. Shape recognition of irregular objects. In David P. Casasent, editor, *Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling*, Proc. SPIE 2904, pages 25–32, 1996.

[11] Teuvo Kohonen. *Self-Organizing Maps*. Springer-Verlag, Berlin, 1995.

[12] S. Marshall. Review of shape coding techniques. *Image and Vision Computing*, 7(4):281–294, November 1989.

[13] Farzin Mokhtarian and Alan Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):34–43, January 1986.

[14] Theodosios Pavlidis. Algorithms for shape analysis of contours and waveforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(4):301–312, July 1980.

[15] Markus Peura and Jukka Iivarinen. Efficiency of simple shape descriptors. In *3rd International Workshop on Visual Form*, Capri, Italy, May 28–30 1997.

[16] Stephen P. Smith and Anil K. Jain. Chord distributions for shape matching. *Computer Graphics and Image Processing*, 20:259–271, 1982.

[17] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image Processing, Analysis and Machine Vision*. Chapman & Hall Computing, 1993.

[18] Charles T. Zahn and Ralph Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, C-21(3):269–281, March 1972.

[19] Pengfei Zhu and Paul M. Chirlian. On critical point detection of digital shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):737–748, August 1995.