A Survey of Shape Analysis Techniques

Sven Loncaric

Department of Electronic Systems and Information Processing

Faculty of Electrical Engineering and Computing, University of Zagreb

Abbreviated title: A Survey of Shape Analysis Techniques

Correspondence address: Sven Loncaric

Department of Electronic Systems and Information Processing

Faculty of Electrical Engineering and Computing, University of Zagreb

Unska 3, 10000 Zagreb, Croatia

Phone: +385-1-612-9891, Fax: +385-1-612-9652

E-mail: sven.loncaric@fer.hr

A Survey of Shape Analysis Techniques

Abstract

This paper provides a review of shape analysis methods. Shape analysis methods play an important role in systems for object recognition, matching, registration, and analysis. Research in shape analysis has been motivated, in part, by studies of human visual form perception systems. Several theories of visual form perception are briefly mentioned. Shape analysis methods are classified into several groups. Classification is determined according to the use of shape boundary or interior, and according to the type of result. An overview of the most representative methods is presented.

Shape analysis   Shape description   Image analysis   Object recognition

1 Introduction

The input to a typical image processing and analysis system is a gray-scale image of a scene containing the objects of interest. In order to understand the contents of the scene it is necessary to recognize the objects located in the scene. The shape of the object is a binary image representing the extent of the object. The shape can be thought of as a silhouette of the object (e.g. obtained by illuminating the object by an infinitely distant light source). There are many imaging applications where image analysis can be reduced to the analysis of shapes, (e.g. organs, cells, machine parts, characters).

Shape analysis methods analyze the objects in a scene. In this paper, we concentrate on shape representation and description aspects of shape analysis. Shape representation methods result in a non-numeric representation of the original shape (e.g. a graph) so that the *important* characteristics of the shape are preserved. The word important in the above sentence typically has different meanings for different applications. Shape description refers to the methods that result in a numeric descriptor of the shape and is a step subsequent to shape representation. A shape description method generates a *shape descriptor vector* (also called a feature vector) from a given shape. The goal of description is to *uniquely* characterize the shape using its shape descriptor vector. The required properties of a shape description scheme are invariance to translation, scale, and rotation. This is required because these three transformations, by definition, do not change the shape of the object.

The input to shape analysis algorithms are shapes (i.e. binary images). The procedures (e.g. image segmentation) used to obtain a shape from a given gray-scale image are not discussed in this paper.

Shape matching or discrimination refers to methods for comparing shapes. It is used in model-based object recognition where a set of known model objects is compared to an unknown object detected in the image. For this purpose a shape description scheme is used to determine the shape descriptor vector for each model shape and unknown shape in the scene. The unknown shape is matched to one of the model shapes by comparing the shape descriptor vectors using a metric.

The problem of the shape analysis has been pursued by many authors thus resulting in a great amount of research. A number of review papers [1, 2, 3], as well as books [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] have been written on the subject of shape analysis.

1.1 Classifications

Shape analysis methods can be classified according to many different criteria. Pavlidis [1] has proposed the following classifications. The first classification is based on the use of shape boundary points as opposed to the interior of the shape. The two resulting classes of algorithms are known as *boundary* (also called *external*) and *global* (or *internal*), respectively. Examples of the former class are algorithms which parse the shape boundary [17, 18, 19, 20, 21, 22, 23, 24] and various Fourier transforms of the boundary [25, 26, 27, 28, 29]. Examples of global methods include the medial axis (also called symmetric axis) transform (MAT) proposed by Blum and described in [30, 31, 32, 17, 33, 34, 20, 22], moment based approaches [35, 36, 37, 38], and methods of shape decomposition into other primitive shapes [39, 40, 41, 42].

Another classification of shape analysis algorithms can be made on the basis of whether the result of the analysis is numeric or non-numeric. For example, the MAT produces another image (containing a symmetric axis) and is therefore called a *space-domain* technique. On the other hand, *scalar transform* techniques produce numbers (scalars or vectors) as results. Examples of later methods include various Fourier [25, 26, 27, 28, 29] and moment-based [35, 36, 37, 38] procedures for shape analysis.

A third classification of shape analysis methods can be made on the basis of information preservation. Methods which allow for the accurate (or at least sufficiently accurate) reconstruction of a shape from its descriptor are called *information preserving* methods, as opposed to methods only capable of partial reconstruction which are called *information non-preserving* techniques. An example of an information non-preserving method is area to perimeter square ratio. Many significantly different shapes can have the same area to perimeter square ratio, and therefore it is not possible to reconstruct the original shape knowing only its area to perimeter square ratio. Many simple shape descriptors suffer from the same problem.

1.2 Shape Description Method Evaluation Criteria

A common problem in shape description research is how to judge the quality of a shape description method. Not all methods are appropriate for every kind of shape and application, i.e. the method of choice depends on the properties of the shape to be described and the particular application. The presence of noise can also influence the choice of method. Consistent evaluation criteria still do not exist for shape description methods. Several authors have proposed evaluation criteria in the form of lists of qualities that a good shape description method should have.

Marr and Nishihara [43] proposed a set of criteria for the evaluation of shape description methods: accessibility, scope and uniqueness, and stability and sensitivity. Accessibility describes how easy (or difficult) it is to compute a shape descriptor in terms of memory requirements and computational time. Scope demonstrates the class of shapes that can be described by the method. Uniqueness describes whether a one-to-one mapping exists between shapes and shape descriptors. Stability and sensitivity describe how sensitive a shape descriptor is to "small" changes in shape.

Brady [44] proposed another set of criteria for the representation of shape.

- *Rich local support.* This criterion requires that a representation is information preserving (rich) and can be computed locally. Local support is important for the representation of occluded objects.

- *Smooth extension* and *subsumption.* This criterion ensures that local descriptions can easily produce global descriptions. This is a kind of continuity of representation.

- *Propagation.* This criterion adds a hierarchical property to representation in the sense that perceptual subparts are propagated from local to global levels of representation.

Brady illustrated these criteria on the Generalized Cylinder representation for three-dimensional objects and the Smoothed Local Symmetries representation for two-dimensional shapes [45].

Both of the methods mentioned above define desired properties in terms of conceptual qualities that cannot be numerically expressed. This limitation renders an exact numerical comparison of shape description methods impossible.

1.3 Organization of Article

Modern shape description research is based both on "classical" methods which completely rely on mathematical and engineering results and on modern research of characteristics of the human visual system. This work attempts to incorporate both elements of shape research since they are closely related, and is organized as follows.

Section 2 emphasizes the role and importance of research in interdisciplinary fields like visual perception, cognition, psychology, and physiology for the development of new shape analysis techniques. Some of the most important shape analysis methods are briefly described in the following sections. These methods are divided into four groups with respect to the use of shape boundary or interior and result type. Sections 3 and 4 present boundary scalar and boundary space-domain techniques, respectively. Sections 5 and 6 address the most characteristic global scalar and global space-domain methods, respectively.

At the end of the article there is a bibliography which the reader may find useful to further explore the field. It is, however, by no means an exhaustive, but intended to serve as a starting point and direct the reader to characteristic research in this area.

2 Human Perception of Visual Form

The human visual system (HVS) is one of the most sofisticated and versatile in nature. Its ability to understand the organization of surrounding nature is unsurpassed by artificially created reasoning systems. There has been a great amount of research activity concentrated on the study of the HVS. One of the purposes of these efforts is to provide a model for developing artificial systems for visual perception and cognition. Early research in the area of artificial systems for shape analysis was isolated from the area of human visual perception, as opposed to more recent research which relies more on results from the study of the HVS and perception.

From the broad field of cognitive science, the areas of visual cognition and perception are of particular interest for the study of shape description. If the structure of the human shape analysis system were known, it would be possible to develop analog artificial systems. For this reason the study of shape analysis methods is often motivated by and utilizes the results of research in the area of human visual perception. An exhaustive survey of human visual perception research is beyond the scope of this paper. Some introductory and more advanced books and articles dealing with visual perception and cognition include [4, 46, 47, 48, 49, 50, 51, 52]. In this section, a brief overview of visual perception research related to shape description is presented.

2.1 Classical Theories of Visual Perception

Several schools of psychology have endeavored to understand and describe the mechanisms of behavior, in general, and the specific aspect of visual perception. Hake [53] discussed several approaches to the representation of natural forms. Zusne [47] presented an overview of contemporary theories of visual form. An in-depth discussion of visual perception theories in psychology is outside the scope of this work, therefore we limit ourselves to a brief discussion of selected topics.

The Gestalt school of psychology [54] has played a revolutionary role with its novel approach to visual form. A detailed exposition of the subject of Gestalt theory can be found in [55, 56, 57, 58]. The Gestalt theory is a non-computational theory of visual form, and thus a disadvantage for practical engineering applications. However, according to Zusne "it is still the only theory to deal with form in a comprehensive fashion" ([47], p. 108). There have been many books on Gestalt laws presenting various lists of principles.

These lists range from six to more than one hundred. Here, we provide a list of laws for visual forms as proposed by Zusne [47].

- Visual form is the most important property of a configuration.
- Visual form is either dynamic or the outcome of dynamic processes which underlie them.
- All visual forms possess at least two aspects, a figured portion called figure and a background called ground.
- Visual forms may possess one or several centers of gravity about which the form is organized.
- Visual forms are transposable (with respect to translation, size, orientation, and color) without loss of identity.
- Visual forms resist change. They tend to maintain their structure against disturbing forces.
- Visual forms will always be as good (regular, symmetric, simple, uniform, exhibiting the minimum amount of stress) as the conditions (pattern stimulus) allow.
- Forms may fuse to produce new ones.
- A change in one part of form affects other parts of the form (law of compensation).
- Visual forms tend to appear and disappear as wholes.
- Visual forms leave an aftereffect that make them easier to remember (law of reproduction).
- Space is anisotropic, it has different properties in different directions.

Another approach to the theory of visual form is found in Hebb's work. Hebb presented a neuropsychological theory of behavior in his book *The Organization of Behavior* [59]. In his theory, Hebb emphasized the role of neural structures in the mechanism of visual perception. His work influenced a number of researchers in the field of artificial neural networks. As opposed to the Gestalt school, Hebb argues that form is not perceived as a whole but consists of parts. The organization and mutual spatial relation of parts must be learned for successful recognition. This learning aspect of perception is the central point in Hebb's theory.

Eye movement is the main mechanism which integrates simpler elements of perception. The simpler elements are angles and lines. Hebb also introduced the notion of *cell assemblies*, which are neurons grouped together so that repeated firings lower synaptic resistance, thus causing neurons in the group to mutually excite each other. Hebb's theory was mostly qualitative and not computational, thus presenting a disadvantage for practical engineering applications.

Gibson [60] developed another comprehensive theory of visual perception. The first principle of his theory is that space is not a geometric or abstract entity, but a real *visual* one characterized by the forms that are in it. Gibson's theory is centered around perceiving real three-dimensional objects, not their two-dimensional projections. The second principle is that a real world stimulus exists behind any simple or complex visual perception. This stimulus is in the form of a gradient which is a property of the surface. Examples of physical gradients are the change in size of texture elements (depth dimension), degree of convergence of parallel edges (perspective), hue and saturation of colors, and shading. Gibson points out that the Gestalt school has been occupied with the study of two-dimensional projections of the three-dimensional world and that its dynamism is no more than the ambiguity of the interpretation of projected images. In his classification there are ten different kinds of form.

- Solid form. (Seeing an object means seeing a solid form.)
- Surface form. (Slanted and forms with edges.)
- Outline form. (A drawing of edges of a solid form.)
- Pictorial form. (Representations which are drawn, photographs, paintings, etc.)
- Plan form. (A drawing of edges of a surface projected on a flat surface.)
- Perspective form. (A perspective drawing of a form.)
- Nonsense form. (Drawings which do not represent a real object.)
- Plane geometric form. (An abstract geometric form not derived from or attempting to make a solid form visible.)
- Solid geometric form. (An abstract part of a three-dimensional space bounded with imaginary surfaces.)

- Projected form. (A plane geometric form which is a projection of a form.)

These forms are grouped into three classes as follows.

- Real objects: solid and surface forms.
- Representations of real objects: outline, pictorial, plan, perspective, and nonsense forms.
- Abstract (non-real) objects: Plane geometric forms, solid geometric forms, and projected forms.

The first class is the "real" class consisting of objects from the real world. The second class are representations of real objects. The third class are abstractions that can be represented using symbols but do not correspond to real objects (because they have no corresponding stimulus in the real world).

A tutorial on visual cognition with an emphasis on shape recognition was written by Pinker [61].

2.2 Modern Theories of Visual Perception

Marr et al. [62, 63, 43, 64, 65, 66] made significant contributions to the study of the human visual perception system. In Marr's paradigm [67], the focus of research is shifted from applications to topics corresponding to modules of the human visual system. An illustration of this point is the so-called *shape from x* research which represents an important part of the total research in computer vision [3]. Papers dealing with shape from x techniques include: shape from shading [68], shape from contour [69, 70], shape from texture [60], shape from stereo [71], and shape from fractal geometry [72].

In [63] Marr developed a primal sketch paradigm for early processing of visual information. In his method, a set of masks is used to measure discontinuities in first and second derivatives of the original image. This information is then processed by subsequent procedures to create a *primal sketch* of the scene. The primal sketch contains locations of edges in the image and is used by subsequent stages of the shape analysis procedure. Marr and Hildreth [65] further developed the concept of the primal sketch and proposed a new edge detection filter based on the zero crossings of the Laplacian of the two-dimensional Gaussian distribution function. In this approach, zeros of Laplacian indicate the inflection point in the edge to detect edge positions.

Koenderink [73] and Koenderink and van Doorn [74] have studied the psychological aspects of visual perception and proposed several interesting paradigms. Conventional approaches to shape are often static in the sense that they treat all shape details equally as global shape features [74]. A dynamic shape model was developed where visual perception is performed on several scales of resolution. Such notions of order and relatedness are present in visual psychology and absent in conventional geometrical theories of shape. It has been argued in [74] that manuals of art theory (such as [75]) exist which have not been given the attention they deserve and which contain practical knowledge accumulated over centuries. In art as well as in perception, a shape is viewed as a hierarchical structure. A procedure for morphogenesis based on multiple levels of resolution has been developed [74]. Any shape can be embedded in a "morphogenetic sequence" based on the solution of the partial differential equation that describes the evolution of the shape through multiple resolutions.

Many authors agree on the significance of high curvature points for visual perception. Attneave and Arnoult [76, 77] performed psychological experiments to investigate the significance of corners for perception. In the famous Attneave's cat experiment a drawing of a cat was used to locate points of high curvature which were then connected to create a simplified drawing of the cat. After a brief presentation the cat could be reliably recognized in the drawing. It has been suggested that such points have high information content. Attneave's work has initiated further research on the topic of curve partitioning [78, 79, 80, 81, 82, 83, 84, 85]. To approximate curves by straight lines, high curvature points are the best place to break the lines, thereby the resulting image retaining in the maximal amount of information necessary for successful shape recognition. For the purpose of shape description, corners are used as points of high curvature and the shape can be approximated by a polygon. A variety of algorithms for polygonal approximation of shape [86, 1, 87, 88, 89, 90, 91] have been developed. Davis [92] combined the use of high curvature points and line segment approximations for polygonal shape approximations. Stokely and Shang [93] investigated methods for measurement of the curvature of 3-D surfaces that evolve in many applications (e.g. tomographic medical images).

Hoffman and Richards [94, 95] argue that when the visual system decomposes objects it does so at points of high negative curvature. This agrees with the principle of transversality [96] found in nature. The principle of transversality contends that when two arbitrarily shaped convex objects interpenetrate each other, the meeting point is a boundary point of concave discontinuity of their tangent planes.

Leyton [97] demonstrated the *Symmetry-Curvature* theorem which claims that any curve section that has only one curvature extremum has one and only one symmetric axis which terminates at the extremum itself. (For more information on symmetric axis work see Section 6.1.) This is an important result because it establishes the connection between two important notions in visual perception. In [98], Leyton developed a theory which claims that all shapes are basically circles which changed form as a result of various deformations caused by external forces like pushing, pulling, stretching, etc. Two problems were considered: the first was the inference of the shape history from a single shape, and the second was the inference of shape evolution between two shapes. The concept of symmetry-curvature was used to explain the process that deformed the object. Symmetric axes show the directions along which a deformation process most likely took place. In [97], Leyton proposed a theory of nested structures of control which, he argues, governs the functioning of the human perceptual system. It is a hierarchical system where at each level of control all levels bellow any given level are also included in information processing.

Pentland [99, 100] investigated methods for representation of natural forms by means of fractal geometry [101, 102, 103]. He argued that fractal functions are appropriate for natural shape representation because many physical processes produce fractal surface shapes. This is due to the fact that natural forms repeat whenever possible and non-animal objects have a limited number of possible forms [104, 105]. Most existing schemes for shape representation were developed for engineering purposes and not necessarily to study perception. Fractal representations produce objects which correspond much better to the human model of visual perception and cognition.

Lowe [106] proposed a computer vision system that can recognize three-dimensional objects from unknown viewpoints and single two-dimensional images. The procedure is non-typical and uses three mechanisms of perceptual grouping to determine three-dimensional knowledge about the object as opposed to a standard bottom-up approach. The disadvantage of bottom-up approaches is that they require an extensive amount of information to perform recognition of an object. Instead, the human visual system is able to perform recognition even with very sparse data or partially occluded objects. The conditions that must be satisfied by perceptual grouping operations are the following.

- The viewpoint invariance condition. This means that observed primitive features must remain stable over a range of viewpoints.
- The detection condition. There must be enough information available to avoid accidental mis-interpretations.

The grouping operations used by Lowe are the following. Grouping on the basis of proximity of line end points was used as one viewpoint invariant operation. The second operation was grouping on the basis of parallelism, which is also viewpoint independent. The third operation was grouping based on collinearity. The preprocessing operation consisted of edge detection using Marr's zero crossings in the image convolved with a Laplacian of Gaussian filter. In the next step a line segmentation was performed. Grouping operations on line-segmented data were performed to determine possible locations of objects.

3 Boundary Scalar Transform Techniques

Boundary scalar transform algorithms typically consist of two major steps. In the first step, a one-dimensional function is constructed from the two-dimensional shape boundary. The constructed one-dimensional function is used in the second step to describe the shape of the two-dimensional boundary. Note that, in this approach, the shape is described *indirectly* by means of a one-dimensional characteristic function of the boundary instead of the two-dimensional boundary itself. Techniques used in the second step of this approach are divided into those based on the Fourier transform of the characteristic function and those based on a stochastic process modeling of the characteristic function.

3.1 From 2-D Shape to 1-D Boundary Representation

2-D shape can be represented using a real or complex 1-D function. In this Section we present several possible 1-D boundary representations of the shape that have been used in literature.

Zahn and Roskies [25] and Bennet and McDonald [107] used a tangent angle versus arc length function. The tangent angle at some point is measured relative to the tangent angle at the initial point. The function is also called the turning function and has been used by Arkin et al. [108] for comparing polygonal shapes.

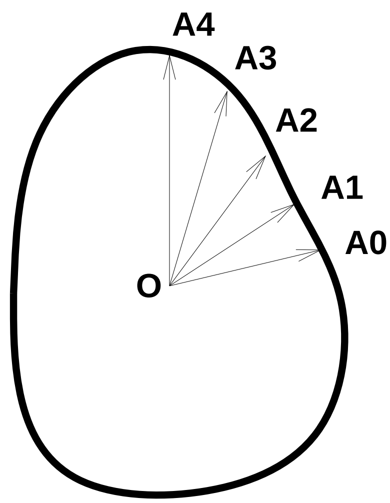

Figure 1: The centroid-to-boundary distance approach.

Granlund [26], Richards and Hemami [27], and Persoon and Fu [28] used a complex function of the form $x(t)+jy(t)$, where t is the arc length parameter. Another approach for the generation of a 1-D representation of the boundary is to use the shape centroid. The shape, its centroid and boundary points are shown in Figure 1. In the first variation of the idea, the values of the 1-D function are equal to distances between shape centroid and boundary points. Boundary points are selected so that the central angles are equal. Another variation of the idea is to use the distance between subsequent boundary points for the 1-D function values. In the third alternative, a shape boundary is approximated with a polygon so that all sides have the same length. The values of the 1-D function are equal to the angle between the polygon sides and a reference line.

Chang, Hwang, and Buehrer [109] constructed the distance function from the centroid to the feature points. In their method, the feature points are the points of high curvature. There are two approaches for detecting high curvature points. The first is to compute curvature directly from the boundary curve. The second approach is to perform a polygonal approximation of the shape and use the polygon vertices as feature points. This is based on the fact that corners are points of high curvature. The computed distances are saved in a linked list. The distances are ordered to achieve rotation invariance. The distances are also divided by the minimal distance to achieve scale invariance. Translation invariance is automatically achieved through the use of centroid.

Wang et al. [110] used a sequence of line segment moments as a 1-D function. Line segments are obtained by partitioning the radial line from the center of the mass to the boundary point. Segments are partitioned into parts within the shape and parts outside the shape.

3.2 Fourier Transform of Boundary

The shape description methods under this category use the Fourier Transform of the 1-D boundary representation to characterize the shape.

The Zahn and Roskies method [25] uses the tangent angle vs. arc length shape boundary representation. The boundary is parsed in a clockwise direction producing negative angles relative to the initial point. The Fourier transform is then applied to the boundary function and the resulting coefficients are used for shape description. Due to the arc length normalization, the shape descriptor is invariant to scale change. The shape descriptor is invariant to translation because the tangent angle function is invariant to shape position. Rotation of the object (i.e. variation of the starting point) causes phase change in the resulting Fourier transform, therefore looking at the magnitude of the Fourier coefficients will ensure rotation invariance of the method. The major advantage of this method is that it is easy to implement and based on a well developed theory of Fourier analysis. The disadvantage is that Fourier transform does not provide local shape information. After the Fourier transform, local shape information is distributed to all coefficients and not localized in the frequency domain. Tangent angle versus arc length representations suffer from very high noise sensitivity because it is difficult to determine the tangent angle for noisy contours.

Pinkowski [111] used Fourier descriptors for the description of shapes appearing in speech spectrograms. This method is used for classifying words containing English semi-vowels. Experiments demonstrated a high recognition rate.

3.3 Bending Energy

Young, Walker, and Bowie [18] proposed an interesting concept of bending energy. According to this approach, a shape can be represented by its bending energy defined by

$$E = \frac{1}{P} \int_0^P |K(p)|^2 dp \tag{1}$$

where $K(p)$ is the curvature function, p is the arc length parameter, and P is the total curve length. To actually compute the bending energy Equation 1 was not used directly, but instead the Fourier transform of the boundary was computed first. Using Fourier coefficients and Parseval's relation, the bending energy was computed in a more efficient way. In addition, the authors proved that the circle was the shape having the minimum average bending energy.

3.4 Stochastic Methods

Methods in this class are based on the stochastic modeling of a 1-D function obtained from the shape as described in Section 3.1. The 1-D function is interpreted as a stochastic process realization. The model parameters obtained by estimation are used as shape descriptors. On the terminology side, the notion of *time-series* is often used in stochastic processes for signals that depend on time. However, it is the 1-D boundary function that is modeled in stochastic shape boundary analysis instead of time function. Regardless of this fact, the term "time-series modeling" can also be found in literature in reference to shape boundary modeling.

Kashyap and Chellappa [19] proposed the use of circular autoregressive models (CAR) for representation of the centroid to boundary points distance function. The CAR model is characterized by a set of unknown parameters and an independent noise sequence. Since the boundary is closed, boundary 1-D function r_t is a periodic function. The particular CAR model that was utilized is the same one that was used by Huang [112]. It is a stochastic process defined by the following m-th order difference equation

$$r_t = \alpha + \sum_{j=1}^{m} \theta_j r_{t-j} + \sqrt{\beta}\omega_t \tag{2}$$

where ω_i are independent random noise sources. The parameters $\{\alpha, \theta_1, \ldots, \theta_m, \beta\}$ are unknown and need to be estimated. The maximum likelihood (ML) parameter estimation [113, 114, 115] was used. The ML estimated parameters θ_i are translation, rotation, and scale invariant. Note that the rotation invariance holds only for angles that are multiples of $2\pi/N$. Parameters α and β are not scale invariant, but the quotient $\alpha/\sqrt{\beta}$ is. Therefore, the vector $[\theta_1, \ldots, \theta_m, \alpha/\sqrt{\beta}]^T$ is used as a shape descriptor. Kashyap and Chellappa further investigated coding and reconstruction schemes and showed that stochastic methods could be used for the description of closed boundaries.

Dubois and Glanz [116] used the same autoregressive model as in [19] but investigated three additional methods for improving pattern classification (shape matching) performance. The classification was then performed by computing the weighed Euclidean distance between unknown object descriptors and training objects. The first improvement was the weighing of the descriptor vector so that components that were common within a training class were emphasized while components that differed were de-emphasized. It has been shown that the optimal feature weight is inversely proportional to the standard deviation of the feature of the class training set [117]. The second improvement consisted in the rotation of the coordinate system and scaling the rotated axes [117]. This groups the members of one pattern class closer in the new coordinate system. The third modification included the use of hyper-planes to divide the pattern space. The least mean squared error procedure [118] yields the optimal hyper-plane parameters. The experimental result showed that the normalized AR model parameters useful shape descriptors.

A modification of the Dubois and Glanz method included statistical knowledge about the boundary noise resulting from the imaging process and boundary sampling [119].

Das, Paulik, and Loh [120] developed a bivariate technique for autoregressive modeling of the shape boundary. They obtained even better experimental results using their technique.

The linear AR model has been compared to the non-linear model of the quadratic Volterra type given by

$$r_t = \sum_{j=1}^{m} a_j r_{t-j} + \sum_{u=0}^{p} \sum_{v=0}^{q} g_{uv} r_{t-u} r_{t-v} + e_t \tag{3}$$

where r_t is the centroid to boundary points distance function [23]. According to Kartikeyan and Sarkar, the linear AR models may not be sufficient for recognition of more complicated (non-convex) shapes. For improved accuracy, a higher order linear model is necessary to increase the dimension of the shape descriptor vector. The use of over fitted AR models may lead to poor recognition performance. However, non-linear models can provide the higher accuracy necessary to describe more complicated shapes. These experiments demonstrated that the quadratic Volterra models performed better classification than the linear AR model [23].

The disadvantage of AR boundary modeling is that in the case of complex boundaries a small number of AR parameters is not sufficient for description. For this reason, He and Kundu [121] combined the use of the AR model with the hidden Markov model. The shape boundary was partitioned into a number of segments

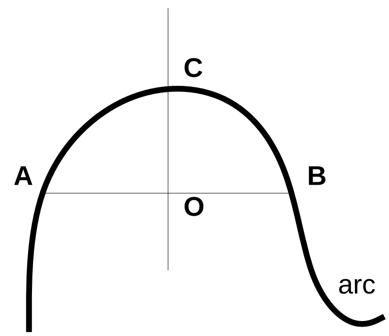

Figure 2: The arc height concept.

and each segment was described using AR modeling. The obtained vectors were analyzed using the hidden Markov model.

3.5 Arc Height Method

Lin, Dou, and Wang [122] used the Arc Height Function (AHF) to characterize the shape boundary. The arc height concept is shown in Figure 2 and defined as follows. An arc chord $\overline{AB}$ of predefined length is positioned on the boundary. The symmetric axis passing through $\overline{OC}$ is drawn perpendicular to the arc chord $\overline{AB}$. The length of the line segment $\overline{OC}$ is called the arc height at position A. As the arc chord is moved along the curve, a mapping between arc length and arc height defines the AHF. The AHF is then used to characterize the shape.

4 Boundary Space Domain Techniques

Boundary space domain methods take shape boundary as input and produce the result in pictorial or graph form. Space domain techniques often appear in various structural approaches to shape recognition. The reason for this is that in structural methods a recognition system processes the visual information in stages starting from the early preprocessing phase to higher levels where the final interpretation of the visual scene is performed. A characteristic of these processing stages is that they produce an image, a graph, or other non-scalar values as opposed to approaches described in Section 3 which produce scalar results. This is why such methods are called space domain methods. Some examples of space domain techniques are presented

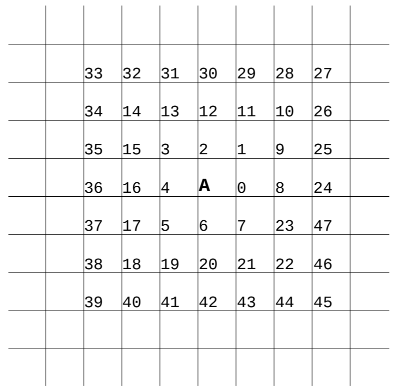

Figure 3: The generalized chain code.

in the following subsections.

4.1 Chain Code

Freeman [123] has proposed a method for coding line drawings called *chain coding.* A more detailed overview of chain code methods and algorithms by Freeman can be found in [124]. The generalized chain code [125] is shown in Figure 3. The nodes surrounding node A are enumerated counter-clockwise in ascending order from inside out. The *link* a_i is a directed line segment. A *chain* is an ordered sequence of links written in the form $A = a_1 a_2 a_3 \ldots a_n$. In addition to these basic operations it is possible to compute the inverse and length of a chain, the integral of a function specified by its chain code, the first and second moments about x and y axes, and the distance between two points connected by a chain. The above operations illustrate the flexibility and versatility of the chain code for algorithm realization.

In [126], Freeman used a chain code description of the boundary to extract the critical points which he then used to generate a shape description that is invariant to translation, rotation, and scale. He also presented another chain-code approach based on the centroidal profile of the shape. The centroidal profile is a plot of the centroid to boundary points distance.

Parui and Majumder [127] used a modified chain code to perform a symmetry analysis. The shape boundary is represented in a hierarchical way so that at the highest level a lower number of polygon vertices is used, while at the lowest level the finest polygonal approximation is utilized. The search for the symmetric axis position is performed by starting at the highest level and shifting to lower levels as the position of the symmetry axis becomes more accurately determined.

4.2 Syntactic techniques

Structural descriptions may be viewed as graphs and as such are suitable for formulation in terms of formal languages [128, 129, 130, 131, 132]. Syntactic methods have several important advantages. They attempt to emulate the structural and hierarchical nature of the HVS. Another advantage is that the theory of formal languages, which syntactic methods rely on, is a well developed field. The main disadvantage of this approach is that shape (or shape boundary) must first be encoded to provide input to the parser. Typically, a low level segmentation of the image must be performed to extract different types of line and curve segments and corners [133]. The obtained boundary features are then formed in a string $S = s_1, s_2, \ldots, s_n$. String element s_i can represent different entities like a chain-code element, a side of polygonal approximation, or an arc. The string of feature symbols is then parsed according to a grammar to detect the shape of the object. In addition to deterministic (non-stochastic) grammars, stochastic grammars have also been investigated [128]. Fu [134] presented the method for image modeling based on stochastic grammars. Attributed grammars manipulate attributes (semantics) at the time of parsing the grammar (syntax). Attributes are usually primitive shape features.

The theory of formal languages, established by Noam Chomsky, has been used in many fields including compiler design, automata theory, computer languages, pattern recognition, and image processing. First, the basic terminology of formal languages is introduced [135]. An *alphabet* is a set of words (symbols). Words are combined together to form a *sentence*. A *language* is a set of sentences that can be composed using the words from the alphabet. Formal languages are defined using grammars. Grammars are sets of syntax rules describing how sentences can be generated using an available vocabulary. Grammar G is the quartet $G = (N, \Sigma, P, S)$, where N denotes the set of nonterminal symbols, Σ is the set of terminal symbols, P is the set of production rules, and $S \in N$ is the start symbol. The language $L(G)$ is a set of sentences generated by the grammar G. The sentences have the following properties.

1. Each sentence is composed of terminal symbols only.

2. Each sentence can be derived from S using an appropriate sequence of production rules from P.

Formal grammars are divided into four classes. Type 0 grammars have the unrestricted form of production rules. Type 1 are context-sensitive grammars where productions depend on the context. Type 2 are context-free grammars and type 3 grammars are regular grammars. Context-free and regular grammars have been used most in practice [131].

4.3 Boundary Approximations

The two most popular schemes for curve approximation are polygonal and spline approximations.

Polygonal approximations are used to approximate the shape boundary using the polygonal line. These methods are based on the use of the minimal error, the minimal polygon perimeter, the maximal internal polygon area, or the minimal external polygon area as approximation criteria. The error measures that are used most are maximal error (yielding various minimax methods) and integral square error.

One of the most popular methods in this group is the *split-and-merge* algorithm [136]. In this approach, a curve is split into segments until some acceptable error is obtained. At the same time, split polygonal segments are merged if the resulting segment approximates the curve within some maximum error. Pavlidis [86] used partial derivatives of the integral square error function to direct Newton's method in search of optimal breakpoints.

Wu and Leou [91] suggested a different optimization criteria to derive polygonal approximations. The internal maximum area, the external minimum area, or the minimum area deviation polygonal approximations were used in their work.

Bengtsson and Eklundh [137] presented a hierarchical method where the shape boundary is represented by a polygonal approximation. The split-and-merge algorithm was used to create the polygonal approximation. The scale-space approach [138] was used to track the position of inflection points on the boundary curve. Stable shape features are those which remain unchanged over scale. Tangents at border points are estimated using a polynomial approximation to yield a multi-scale representation of the curve.

Splines have been very popular for the interpolation of functions and the approximation of curves. Ikebe and Miyamoto [139] wrote an overview of spline applications for shape design, representation, and restoration. Mathematical treatment of splines is presented in several books [140, 141, 142], while the computer graphics perspective is presented in [143, 144, 145, 146].

Splines posses the beneficial property of minimizing curvature. In other words, they approximate a given function with a curve having the minimum average curvature. This makes them perfect candidates for the "natural" representation of curves. The disadvantage of splines in interpolation problems is that the local function value modification changes the complete spline representation. B-splines are constructed so that the local function value change does not spread to the rest of the intervals. B-splines can be used for the interpolation of plane curves given by parametric equations $(x(t), y(t))$. In this case, each parametric equation is interpolated independently. Cohen et al. [147] proposed a technique for curve representation and matching using B-splines.

Surfaces are often represented by means of a family of curves. The simplest solution is just a pointwise linear interpolation between curves. This technique is called *lofting* and is used widely in shipbuilding and aircraft industries [139].

Chung et al. [148] developed a method for the polygonal approximation of a shape by means of the Hopfield neural network. The approximation is formulated as a minimization of the network energy function which is defined as the arc-to-chord deviation between the curve and the polygon.

4.4 Scale-Space Techniques

This group contains methods that rely on the scale-space representation. Witkin [138] proposed a scale-space filtering approach which provides a useful representation for representing significant object features. The representation was created by tracking the position of inflection points in signals filtered by low-pass Gaussian filters of variable widths. The inflection points that remained present in the representation were expected to be "significant" object characteristics.

Babaud et al. [149] proved the uniqueness of the Gaussian kernel for scale-space filtering. The Gaussian kernel has the desirable property of saving inflection points when the width of the filter is increased. In other words, contours in scale-space image cannot disappear when the filter width is increased. The Gaussian filter is the only filter with such property.

Asada and Brady [150] proposed a new approach based on the ideas developed by Marr et al. to introduce a representation called the *curvature primal sketch.* This is a scale-space approach for the representation of curvature and is similar to Marr's primal sketch for edge detection. The shape boundary is filtered with Gaussian functions of increasing width to obtain a multi-scale representation of shape boundary. The curvature is then computed at different scales to obtain the curvature primal sketch.

Mokhtarian and Mackworth [151, 152] applied the scale-space approach to the description of planar shapes using the shape boundary. The curvature along the contour was next computed and smoothed with variable width Gaussian filters. The scale space image of the curvature function was used as a hierarchical shape descriptor that is invariant to translation, scale, and rotation.

The concept of multi-scale filtering is also present in mathematical morphology. Chen and Yan [153] used a variable size structuring element to perform various morphological operations.

Dill et al. [154] studied the role of leukocyte locomotion. They created multiple smoothed versions of a leukocyte boundary to extract skeletons. The boundary was represented using the chain 1-code. A Gaussian filter of variable width was used to create smoothed versions of the boundary. The skeleton was computed using a technique based on Arcelli's algorithm [155]. This algorithm improves noise sensitivity and sensitivity to global convexities.

4.5 Boundary Decomposition

The methods in this group decompose the shape boundary into curve segments. H. Liu and M. Srinath [156] developed a procedure for shape classification using contour matching. The input to the procedure was an object shape. The Sobel edge detector was used to compute the direction gradients and the tangent angle along the boundary. The tangent angle function was convolved with the derivative of the Gaussian function to find the smoothed curvature function. The boundary was segmented at points of high curvature. Curve matching was performed in two steps. In the first step, individual segments for two shapes were compared. In the second step, groups of segments were compared and a group was disqualified if less then three consecutive segments matched. The final step was used to measure the distance between two shapes. This was done by using the Chamfer 3/4 distance transform [157, 158] which approximates the Euclidean distance transform very well, but is less computationally intensive. The Chamfer 3/4 distance transform was computed for the first boundary; the second boundary was superimposed on the distance transform of the first boundary and the average distance value was computed. Experimental results demonstrated the feasibility of the method for shape matching.

5 Global Scalar Transform Techniques

The methods classified here compute a scalar result based on the global shape. Moment based methods are among the most popular methods from the group of global (or internal) scalar transform methods. Shape vectors and matrices are among the lesser known methods for shape description.

5.1 Moments

Moments were first used in mechanics for purposes other than shape description. Recent surveys of the field were written by Prokop and Reeves [38], and Weiss [159], who takes a more general approach.

The two-dimensional Cartesian moment m_{pq} of order $p+q$ of the function $f(x,y)$ is defined as

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy \tag{4}$$

The use of moments for shape description was initiated by Hu [35]. He proved that moment-based shape description is information preserving. Moments m_{pq} are uniquely determined by the function $f(x,y)$ and vice versa the moments m_{pq} are sufficient to accurately reconstruct the original function $f(x,y)$. The zeroth order moment m_{00} is equal to the shape area assuming that $f(x,y)$ is a silhouette function with value one within the shape and zero outside the shape. First order moments can be used to compute the coordinates of the center of the mass as $x_c = m_{10}/m_{00}$ and $y_c = m_{01}/m_{00}$. Second order moments are called moments of inertia and can be used to determine the principal axes of the shape. Principal axes are axes with respect to which there are maximum and minimum second order moments. Moments of projections are actually one-dimensional moments of projection functions. The moments defined by Equation 4 are not ideal for shape description since they are not invariant to translation, rotation, and scale. To overcome this difficulty, Hu [35] proposed seven invariant moments (also called moment invariants). These moments do not depend on the position, orientation, or scale of the shape. A generalization of moment transform to other transform

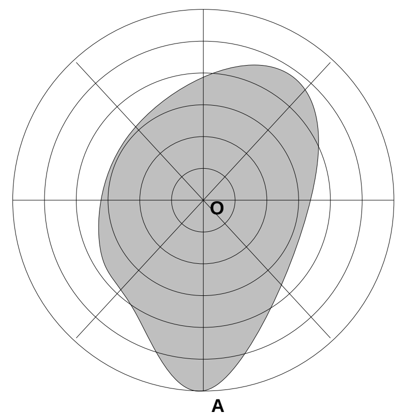

Figure 4: The shape matrix concept.

kernels is possible by replacing a conventional transform kernel $x^p y^q$ by a more general form $P_p(x)P_p(y)$. Particularly appealing is to use orthogonal polynomials as moment transform kernels [36]. In this case, the moments produce minimal information redundancy [38]. This is important for optimal utilization of the information available in a given number of moments. Some of the orthogonal polynomial systems that have been investigated [36] include Legendre, Zernike, and pseudo-Zernike polynomials. The advantage of moment methods is that they are mathematically concise. The disadvantage is that it is difficult to correlate high-order moments with shape features. As with most scalar transform methods, local information and shape features cannot be detected.

An alternative transform approach is the Fourier transform of the shape. The disadvantage is that it is impossible to 'sense' local shape features and it is also computationally intensive. The problem of invariance to translation, scale, and rotation is also present in this approach.

5.2 Shape Matrices and Vectors

Shape matrix and vector approaches use global shape information to create a numerical (matrix or vector) description of the shape.

Goshtasby [21] used a matrix to represent the pixel values corresponding to a polar raster of coordinates centered in the shape center of a mass. This idea is illustrated in Figure 4. A polar raster of concentric circles and radial lines is positioned in the shape center of the mass. Line $\overline{OA}$ represents the axis of the polar coordinate system. The binary value of the shape is sampled at the intersections of the circles and radial lines. The shape matrix is formed so that the circles correspond to the matrix columns and the radial lines correspond to the matrix rows. This scheme is invariant to translation, rotation, and scaling. The maximum radius of the shape is equal to the radius of the circle centered in the shape center of the mass that contains the shape.

In the method used by Taza and Suen [24], shapes were described by means of shape matrices and a comparison of matrices was performed to classify unknown shapes into one of the known classes. A scheme for weighing matrix entries was developed for more objective comparison. The weighing was based on the fact that sampling density is not constant with the polar sampling raster. Without weighing the innermost shape samples are implicitly given much more importance than peripheral shape pixels since sampling density is much higher in the center of the shape.

Parui, Sarma, and Majumder [160] proposed a shape description scheme based on the relative areas of the shape contained in concentric rings located in the shape center of the mass. Let L be the maximum radius of the shape S to be described. Let T_k be the k-th ring of n concentric rings obtained by sectioning the maximum radius L into n equal segments. Note that $S \subset \cup_{i=1}^{n} T_i$. Let

$$x_i = \frac{A(S \cap T_i)}{A(T_i)} \tag{5}$$

where $A(.)$ is the function that returns the area of its argument. In other words, x_i is the area of shape S contained in ring T_i relative to the area of the ring itself. The shape descriptor vector is formed as $x = [x_1 \ldots x_n]^T$. The authors demonstrated that the shape vector scheme can be used for shape matching.

5.3 Morphological Methods

Mathematical morphology has evolved as a useful tool for various image processing tasks [161, 162, 163, 164, 165]. It is suitable for shape-related processing since morphological operations are directly related to the object shape.

It is first necessary to introduce some basic definitions related to mathematical morphology. Morphological operations are defined in terms of set theory. For two dimensional shapes we consider sets which are subsets of R^2, and denoted by capital letters in the following text. The multiplication of a set X by a real number r is defined as

$$rX = \bigcup_{x \in X} \{rx\} \tag{6}$$

The translation of a set X by a real two-dimensional vector y is defined as

$$X_y = \{x + y \mid x \in X\} \tag{7}$$

The symmetric set $\check{X}$ of a set X is defined as

$$\check{X} = \{x \in R^2 \mid -x \in X\} \tag{8}$$

Minkowski addition and subtraction are defined by

$$X \oplus Y = \bigcup_{y \in Y} X_y \tag{9}$$

$$X \ominus Y = \bigcap_{y \in Y} X_y \tag{10}$$

Morphological erosion and dilation operations are the basic and among the most useful operations for image processing purposes. Most other operations are derived from these two. Morphological dilation and erosion are defined by

$$\mathcal{D}(X, B) = X \oplus \check{B} = \bigcup_{y \in B} X_{-y} \tag{11}$$

$$\mathcal{E}(X, B) = X \ominus \check{B} = \bigcap_{y \in B} X_{-y} \tag{12}$$

Morphological opening and closing are among the most powerful operations in mathematical morphology. Morphological opening and closing are defined as

$$X \circ B = X \ominus \check{B} \oplus B \tag{13}$$

$$X \bullet B = X \oplus \check{B} \ominus B \tag{14}$$

Set B is often called a structuring element because its shape determines the structure of shape X that will be affected by morphological processing.

A number of approaches to shape description based on mathematical morphology have been investigated. In this section, shape description methods based on the area of morphologically processed images are presented.

Maragos [166] proposed the concept of pattern spectrum. Pattern spectrum for continuous images is defined by

$$PS_X(r, B) = \frac{-dA(X \circ rB)}{dr} \tag{15}$$

where B is a unit disk structuring element and $A(X)$ denotes the area of set X. The pattern spectrum of set X is obtained by opening X with a disk of variable size. The areas of the resulting sets were measured. The pattern spectrum is defined as the derivative of the area function with respect to r, the radius of the disk structuring element. In addition to the continuous case, Maragos proposed a pattern spectrum definition for discrete images and gray scale morphology.

The pattern spectrum approach is related to the notion of granulometries first studied by Matheron [161] and more recently by Dougherty [167, 168]. Granulometries are a result of a sieving operation applied to binary images. It is a sieving operation because the structure in the image is filtered according to component (or particle) size. The result of sieving is a sequence obtained by opening the shape by a sequence of structuring elements. The sequence of openings is called a granulometry. The decreasing sequence of areas of successive openings is called a size distribution. Note that the negative derivative (or difference in the discrete case) of the area sequence is equal to the Maragos pattern spectrum.

Shih and Pu [169] proposed another "spectrum" transformation called the G-spectrum, which is an extension of the work of Goutsias and Schonfeld (see Section 6.2). The authors proved that their representation is less redundant than granulometric size distributions or pattern spectrum.

Another set of techniques were derived from the concept of morphological covariance [162]. Loui et al. [170] used the geometrical correlation function (GCF) for representation of two-dimensional shapes. The GCF is based on morphological covariance [162] The authors used the GCF for shape description and matching. This method has rotation and translation invariance. If scale invariance is desired, a preprocessing step of rescaling can be added. Rotation invariance was achieved through the use of minimum or maximum correlation functions. Experimental results have shown that this method is useful for shape matching.

Maragos [171, 172] related the mean absolute error criterion for signal matching to the morphological correlation function. The morphological cross-correlation function can be related to the minimum absolute error matching as follows. Since $|a - b| = a + b - 2\min(a, b)$, the minimization of absolute error is equivalent to the maximization of $\min(a, b)$. A similar argument has been used in the past for minimization of the mean square error. In that case, $(a - b)^2 = a^2 + b^2 - 2ab$ and minimization of the mean square error leads to a maximization of the conventional linear cross-correlation function between signals.

Shapiro et al. [173] used the residual approach to shape matching. The algorithm used one resolution and a single structuring element shape (disk). The residual image centroid, area, and ratio of minor to major axis of the best fitting ellipse were used to represent the shape.

Loncaric and Dhawan [174] developed a method for shape description called Morphological Signature Transform (MST). The MST method for shape description utilizes multi-resolution morphological image processing by multiple structuring elements (SEs). This method attempts to combine multi-resolution image processing [175] with the power of mathematical morphology. The MST shape representation method is based on the decomposition of a *complex* shape to multiple *simple* signature shapes. The idea of this approach is to process decomposed, multiple shapes instead of the original shape. The decomposed shapes are called *signature* shapes because they contain information about the original shape which is extracted through the property decomposition process.

The MST decomposition of shape X with respect to a (not necessarily convex) structuring element S is defined as: $X_S(r, n) = \mathcal{M}(rX, S^n)$, where $r \in R$, and $n \in N$. $\mathcal{M}$ is a binary morphological operator (e.g. erosion, dilation, opening, and closing). A short notation for S^n is defined as $S^n = S \oplus S \oplus \cdots \oplus S$, where there are n summands in the Minkowski addition on the right side of the equation. Sets $X_S(r, n)$ are also called shape signatures. The MST shape description uses the areas of the shape signatures obtained using multiple SEs and multiple object scales to generate shape descriptors. Multiple SEs are obtained by rotating single or multiple original structuring elements.

A method for near-optimal shape matching using MST was developed in [176]. It is based on a genetic algorithm [177, 178, 179] for selection of a near-optimal structuring element for MST shape description. The selected SE provides nearly the best discrimination of shapes from a given class.

6 Global Space Domain Techniques

Global space domain methods are based on the analysis of the global shape. The resulting shape descriptor is non-scalar (e.g. a graph or an image). The most representative methods from this group are discussed in the following sections.

6.1 Medial Axis Transform

The most popular and the most studied global space domain method is the *medial axis* transform (MAT) originally proposed by Blum [30, 32]. The idea of this approach is to represent the shape using a graph and hope that the important shape features are preserved in the graph. On the terminology side, the medial axis was originally used by Blum. A skeletal pair consisting of the skeleton and the "quench" function is used by Calabi [180]. The terms *prairie fire* transform, *symmetric axis* transform, and *skeleton* transform have been used in literature to refer to the same approach. Additional material on this topic can be found in [31, 17, 34, 20, 22].

This approach is motivated by the study of neural physiology and visual psychology. In particular, Blum hypothesized that the process of image formation on the retina is a chain reaction in the following sense. When an object image is formed on the retina a certain number of excited neurons are fired, lowering the excitation levels of neighboring neurons and causing them to fire a short interval of time later (inhibition). This process is repeated until the whole area of the object is "tiled" with firing neurons. The inhibited neurons cannot fire again for a short time due to underlying neurophysiological processes [181]. Therefore the wave front of the firing cells cannot move back towards the retinal areas containing inhibitory neurons. This mechanism is similar to the spreading of a prairie fire. In fact, the first approach Blum used was a temporal function showing the arc length of wave front versus time. This approach did not prove very useful for shape description purposes. Blum's second concept, the concept of medial axis, has proven to be more useful for shape description purposes.

The purpose of the medial axis transform is to extract a skeletal, stick-like figure from the object. This figure can, hopefully, be used to represent and describe the object shape. The formation of the skeleton can be explained by the following example. Let the object interior be composed of a burnable dry grass while

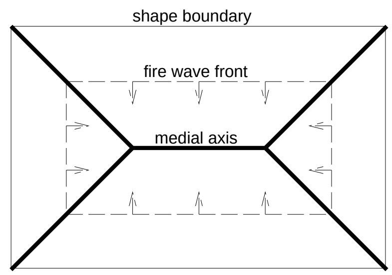

Figure 5: The medial axis transform.

the object exterior is composed of unburnable wet grass. If fire is set simultaneously at all points of the shape boundary, it will propagate towards the center of the shape. However, at some positions the wave front from one direction will intersect the fire wave front coming from another direction, thus extinguishing the fire. Points of wave front collision are called quench points. The skeleton of the figure is defined as the set of quench points. An example is shown in Figure 5. It is possible to reconstruct the shape using its skeleton and quench function. The quench function $q(x)$ at some point x of the skeleton S has a value equal to the radius of the touching circle at that point. The touching circle at some point has the center in that point and touches the boundary at at least two other points. To reconstruct the object one must position a disk of radius $q(x)$ at location x, for each $x \in S$. The union of all disks is equal to the original shape. Wave front propagation can be computed using the distance transform methods [182, 183] or the shrinking operation [31]. Gray scale images have been processed using the min-max operators [184, 34, 185]. The shrinking and related Boolean-nature local neighborhood operators have been used intensively in cellular array image processors [186, 187, 188, 189, 190, 191, 192, 193]. The disadvantage of the medial axis transform is that it is very sensitive to noise on the shape boundary. Even small changes in the shape can cause significant changes in the topology of the MAT graph. Another difficulty is in the practical realization in discrete spaces [194]. The change from continuous to discrete space causes some difficulties. For example, the MAT graph of a connected object may not be connected.

To resolve the problem of noise, Blum and Nagel [33] proposed a *generalized medial axis* transform, based on the touching circle defined as a circle which is tangent to the shape boundary without intersecting it. The *r-symmetric axis* of a shape is defined as the union of all points that have a touching circle of radius greater than r and at least two points that touch the boundary. The requirement that the radius is greater than r

prevents little noisy spikes from generating skeleton segments. The two touching point requirement limits the selection to only skeleton points. The generalized skeleton has proven to have better noise robustness [33]. The concept of symmetry and symmetric axis has been further developed in works by Brooks [195] and Brady [45].

The problem of determining the axis of symmetry is related to an inverse problem - the generation of shapes from their symmetric axis. Objects generated in such a way are called *ribbons*. The synthesis of a ribbon-like object is done by means of the axis and the *generator*. The generator shape is a geometric figure that moves along the axis and changes size as it moves. In Blum's approach, a disk was used as the generator and the constructed shapes are called *Blum ribbons*. *L-ribbons* are generated using the line segment. L-ribbons with a fixed angle between the line segment and the axis are called *Brooks ribbons* [195]. L-ribbons with a fixed angle between the line segment and the side of the ribbon are called *Brady ribbons* [45]. Rosenfeld [196] and Ponce [197] compared various ribbon generating procedures.

In the following text we present several definitions of symmetry that have been proposed in literature. Symmetry is always defined in terms of a condition that has to be satisfied for two points to form a symmetry; the line connecting two points is called the line of symmetry. The middle point is the point in the middle of the line of symmetry. The axis of symmetry is formed by the union of middle points.

Skew symmetry [198] is a symmetry where the lines connecting corresponding (symmetrical) points are perpendicular to the axis of symmetry. This is a mirror-type symmetry. The notion of *parallel* symmetry was proposed by Ulupinar and Nevatia [70]. This symmetry assumes that there is a continuous monotonic mapping between parameters in the parametric representation of symmetric curves so that tangent angles of both curves are the same. Let $c_i(p) = (x_i(p), y_i(p))$ where $i = 1, 2$ be the parametric curve equations and $\phi_i(p)$ be the tangent angle equation. For c_1 and c_2 to be parallel symmetric there must exist a function $f(p)$ such that $c_1(p) = c_2(f(p))$. Brady [45] developed the notion of *smooth local* symmetry. Here, two points form local symmetry if the angles between the curve normals at two given points and the line connecting the points are the same.

A hierarchical (multi-resolution) approach to skeleton analysis was presented by Pizer, Oliver, and Bloomberg [199]. In their approach, hierarchy by scale was used to construct a series of skeletons. This multi-scale approach has been demonstrated to overcome the noise dependencies of the representation [199].

Rom and Medioni [200] proposed a hierarchical representation of shape using axial shape description. This approach combines several paradigms mentioned above. The original shape was broken into parts at negative curvature minima of the shape boundary [94]. The obtained parts were represented using smooth local symmetry ribbons. Parallel symmetry was used to capture the global relationship between parts. Typically, the procedure to determine the axes is computationally expensive and produces many unwanted axes. The unwanted axes can be eliminated using various approaches [201]. In the final stage, a recursive procedure was performed for shape decomposition. In each step of the recursive procedure the axial representation of the shape was made and its smallest parts removed. The shape was then recreated and the procedure repeated. The series of produced shapes represent the decomposition. In [202] B-splines were used to find a boundary approximation from the edge map. The procedures for the computation of skew, parallel, and smooth local symmetries were presented for the B-spline approximated boundary.

Leymarie and Levine [203] developed a new method for the extraction of symmetry axis which does not suffer from the discretization problems that many other algorithms do. This method was based on the use of *snakes* for active contour representation, high curvature points on the boundary, and symmetric axis transform. The result was a dynamic multi-scale skeleton representation. Axes of symmetry were primarily extracted from binary images, but Gauch and Pizer [204] proposed a method for extraction of the *intensity axis* of symmetry from gray scale images. This method was applied to a shape-based image segmentation where it was possible to identify parts of the object corresponding to different components of the intensity axis of symmetry. Han and Fan [205] proposed a method for skeleton extraction using a boundary representation of the shape. In the first step, the shape boundary is extracted using a contour vectorization algorithm. In the second step, pairing of contours is done using a contour matching algorithm. In the final step, the skeleton is found. This method avoids some of the problems typical in conventional skeletonization algorithms. Shih and Pu [206] developed a skeletonization algorithm by tracking the maximum values of Euclidean distance transform. This method has several advantages, such as connectivity preservation, single pixel width, and the resulting skeleton is similar to most other symmetrical axes. Ogniewicz and Kuebler [207] performed the skeletonization of a planar shape based on Euclidean metric and preserves connectivity. This method uses the Voronoi diagram to represent the shape.

Maragos and Schafer [208] used mathematical morphology to extract skeleton subsets for efficient coding of binary images. This method was based on the work of Lantuejoul and Serra, who demonstrated [162] that the skeleton could be represented as a union of components and that the original shape could be reconstructed using these components. The skeleton components were encoded using block and run-length coding methods [112]. The coding was efficient because the skeleton subsets were thin binary images.

6.2 Shape Decomposition

In shape decomposition techniques a shape is represented as a combination of component shapes. The idea is to represent complex shapes in terms of simple components. Pavlidis stated the problem of global shape analysis (decomposition) as follows.

> Among the boundary points find sets of points which are closely related. Such sets may be used to assign labels to corresponding parts of the object [2].

In this approach, shape decomposition is based on the properties of boundary points. Several authors have used this approach for shape decomposition. Decomposition criteria can be defined in terms of the medial-axis transform, and can require convex components or visibility of boundary points. In the medial axis transform approach, two boundary points are labeled related if they are both on the circle contained in the object shape. Decomposition criteria can be formulated to require that the line segment between two points on the boundary is contained in the shape that is described. This kind of criteria leads to decomposition into convex components. Boundary point clustering [209, 210] is a probabilistic method which requires that decomposition is done so that each point in the boundary is visible by most other points. Decomposition on the basis of k-nearest neighbors [211] corresponds to approximating polygon sides instead of boundary points. In stroke detector approaches points are related if they are close to each other across the boundary. Semantic considerations about shapes being described were taken into account in the method for shape decomposition by collinearity [41]. This method was based on heuristics based on human concepts of collinearity. Experiments verified the effectiveness of the approach. Vanderheydt et al. [212] used the fuzzy subset theory [213, 214] to direct decomposition based on convex and concave boundary points and