A Syntactic Approach to Scale-Space-Based Corner Description

Cornelia Fermüller and Walter Kropatsch

Fig. 1. Curvature points of a curve.

Abstract—Planar curves are described by information about corners integrated over various levels of resolution. The detection of corners takes place on a digital representation. To compensate for ambiguities arising from sampling problems due to the discreteness, results about the local behavior of curvature extrema in continuous scale-space are employed.

Index Terms—2-D-shape description, scale-space reduction of curvature extrema, corner detection, multiresolution representation.

I. MOTIVATION AND PREVIOUS RESEARCH

The problem of finding a shape description of planar curves through a representation of points of interest along the curve [6], [13] has received a lot of attention in the past. Previously published methods can be broadly classified into those performing corner detection at one scale (usually through polygonal approximation of the curve [20]) and those dealing with descriptions at different scales [7], [17]. The latter are further classifiable into methods that deal with the problem in a continuous manner, in scale-space [1], [16], and methods that represent the data in a discrete way by employing multiresolution structures (e.g., pyramids [9]) or using symbolic representations of features at multiple scales [21].

Techniques that operate at just one level of resolution may suffer from the disadvantage of finding many unimportant details while at the same time missing large rounded corners. Techniques that operate in scale-space on a continuous representation are quite elaborate, involving considerable overhead; therefore, various discretization schemes have been introduced [2]. Although scale-space methods have produced interesting results, they may be problematic in practical applications, since they must employ either 1-D [16] or 2-D smoothing [8], [19]. In the first case, important large-scale structures may be lost; while in the second case, the topological properties may be destroyed. Finally, techniques that operate on a discrete pyramid, where the number of grid points is reduced from one level to the next, are limited to a finite number of resolutions and may suffer from the problem of undersampling.

The method introduced in this paper, which employs "syntactic smoothing" (to be explained later), works on a discrete pyramid [11], but takes advantage of mathematical relationships among curves in scale-space, and can thus be considered as a hybrid algorithm. Although it is of a discrete nature, it is supported by scale-space information. Furthermore, it combines the advantages of 1-D and 2-
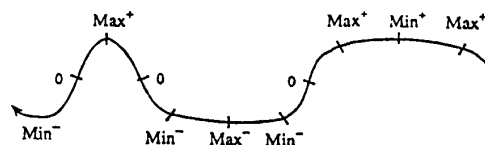
D smoothing, since local 2-D smoothing is performed, but the context information inherent in the curve is considered.

The motivation behind this work that first appeared in [5] is to introduce a curve description which is suitable for many higher-level visual tasks, such as matching used in problems related to stereo, motion, or object recognition. A description of curves should clearly be robust under rotation, scaling, and translation. Further criteria of importance for a reliable computer description are: the computability of the representation by using only local support, the representation of the description at varying levels of detail, and its stability; that is, small changes in the input should cause only small changes in the representation.

The curve pyramid [11] is first used in order to obtain a representation of the curve at varying levels of detail. Different resolutions of curves in digital images are calculated by reducing a small number of curve segments at higher resolution to one segment at lower resolution. The images (the levels of the pyramid) are superimposed on each other in such a way that there exists a geometrical relationship between their elements.

Then, a method for calculating corners in parallel is introduced. It is based on the idea of deciding whether a pixel represents a corner by looking only at the pixel itself and a few of its neighbors. Continuous curves in scale-space are considered to analyze the behavior of curvature extrema under smoothing. The results obtained are used to define measures for the description of a curve in the pyramid. These measures form the basis of a stable description.

II. CURVATURE POINTS OF CONTINUOUS CURVES IN SCALE-SPACE

Planar curves are described by points where the curvature has a local extremum or has the value zero.

Definition 1 (Curvature Points): $C(t)$, with t any parameterization, is an oriented, planar, closed curve. The maxima, minima, and zero-crossings of the curvature are called *curvature points.* Among the extrema, there is a further distinction as to whether the value of the curvature at these points is positive or negative. Therefore, there exist five classes of points: positive maximum (Max^+), negative maximum (Max^-), positive minimum (Min^+), negative minimum (Min^-), and inflection point (0) (Fig. 1).

This classification depends on the orientation of the curve. If it changes, positive maxima become negative minima, negative maxima become positive minima, and vice versa.

Our method of curve description using curvature employs both the size (value) and scope of the curvature.

Definition 2 (Scope of Curvature): Let $C(t)$ be an oriented, planar, closed curve. The *scope of curvature* ($B(t_k)$), with center at point C_{t_k}, consists of the right and left scope $B(t_k) = \{BR(t_k), BL(t_k)\}$. The right scope is the length of the curve's arc from C_{t_k} to the next curvature point in the positive orientation, and $BL(t_k)$ is the
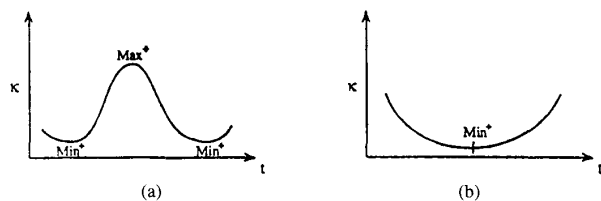
(a)   (b)

Fig. 2. Reduction of curvature points.

Fig. 3. Reduction of curve code in the pyramid.

arc's length from C_{t_k} to the next curvature point in the negative orientation.

In order to provide the means of relating descriptions at different scales, we need to analyze the behavior of curves under progressive smoothing. This method is called scale-space-filtering [22]. It is a way of describing a curve $C(t, \sigma)$ under smoothing with a kernel of width σ, where σ is treated as a continuously increasing parameter. $C(t, \sigma)$ is the convolution ($\star$) of a curve $C(t)$ with a kernel $g(t, \sigma)$:

$$C(t, \sigma) = C(t) \star g(t, \sigma).$$

In principle, there are several possibilities for choosing $g(t, \sigma)$. Babaud *et al.* [3] and Yuille and Poggio [23] proved that when filtering a one-dimensional function $x(t)$ with a Gaussian, no generic zero-crossings and no curvature extrema are created as the scale increases. Bergholm [4] showed that when blurring with a two-dimensional Gaussian, any closed curve turns into a circle.

An equivalent way of generating the family of signals in scale-space is by solving the diffusion equation [10]. Lindeberg [14] has analyzed the nature of smoothing kernels when dealing with discrete signals, which led to the development of a discrete analog of Gaussian kernels and to a discretized version of the diffusion equation.

The previous studies show that the number of maxima, minima, and zero-crossings of curvature, for curves that are smoothed in these ways, decreases. In this paper, curves are characterized by their corners; therefore, the behavior of maxima and minima of curvature under smoothing is analyzed. The following theorem makes this explicit.

Proposition 1: There are ten possible combinations of three successive curvature points, when the middle one is an extremum as listed below. The local reduction of these triplets under smoothing is shown below.

(R1)	Min^+	Max^+	Min^+	$\rightarrow$	Min^+
(R2)	0	Max^+	Min^+	$\rightarrow$	0
(R3)	Min^+	Max^+	0	$\rightarrow$	0
(R4)	0	Max^+	0	$\rightarrow$	Max^-
(R5)	Max^+	Min^+	Max^+	$\rightarrow$	Max^+
(R6)	Min^-	Max^-	Min^-	$\rightarrow$	Min^-
(R7)	Max^-	Min^-	Max^-	$\rightarrow$	Max^-
(R8)	Max^-	Min^-	0	$\rightarrow$	0
(R9)	0	Min^-	Max^-	$\rightarrow$	0
(R10)	0	Min^-	0	$\rightarrow$	Min^+

To prove the validity of each reduction, one needs to plot the curve in 2-D space, with the x-axis representing the arc length and the y-axis the curvature, and observe in this space when the middle extremum disappears [5]. For example, R1 can be established by comparing (a) and (b) of Fig. 2.

At this point, we need to emphasize that the results of the theorem are of a syntactic nature and do not involve the smoothing parameter σ. The value of σ at which the middle extremum in any of the rules (R1–R10) disappears depends on the size and scope of curvature at the points under consideration.

III. DIGITAL REPRESENTATION OF CURVES AND CORNER DETECTION

In this study, the curve code [12] is used for encoding curves. A digital image is overlaid with a grid. Curves are represented by their intersections with the sides of the square grid cells. This information about intersections is stored in the cells.

By merging the contents of the cells, and producing in this way a stack of images of different cell sizes and different resolutions, the curve pyramid is obtained. In this representation, the cells' areas at each resolution are twice as large as those at the next lower resolution. The lattices are rotated by 45° from level to level [12]. In the first step, the squares are divided by a diagonal into two triangles (operation split); and in the second step, groups of four triangles are merged and their contents are reduced to the content of one cell at the lower resolution (operation merge) [11] (Fig. 3).

This syntactic method of reducing the resolution offers an alternative to one- and two-dimensional smoothing. As mentioned earlier, both of these methods have disadvantages: two-dimensional smoothing does not necessarily maintain the topological properties; and one-dimensional scale-space descriptions, such as the analytical one described above, perform less well with spike-like, highly convex or concave features [15]. With the proposed method, these problems can be overcome because the reduction used is a smoothing of curves in both dimensions, which take into account 1-D context information.

In the curve, pyramid corners are detected with a simple, local method. The curve code elements of neighboring picture elements are checked for whether they possibly could originate from a straight line. The method considers just three or five connected code elements. If they cannot be due to a straight line, the sequence of code elements is recognized as originating from curvature, and the central element is recognized as a corner. In this way, two types of curvature points are considered: Max^+ and Min^-. Curvatures whose estimation need more than five pixels will not be detected in this step. We call the procedure using only three elements the *three-element method* and the one using five elements the *five-element method*.

IV. DETECTABILITY OF CORNERS

A. Necessary Corner Conditions

The particular choice of the curve code out is determined by its position in the pyramid. Since the description should be independent of the specific reduction, the *size* of the *angles* in the corner sequences and the *distance* between two neighboring corners are examined and *necessary corner conditions* are developed. These conditions have to be satisfied by corners in order for them to be detectable under all possible reductions.

For the three-element method (five-element method), the straight lines forming a corner must enclose an angle of 63.4° (108.4°), and two neighboring corners have to be at one receptive field (a receptive region of three code-elements) distance. A receptive field of a cell at a level k in the pyramid is defined as the region from which this cell obtains its information; and a receptive region is the union of receptive fields corresponding to neighboring cells. If we detect corners with the three- or five-element method, and exclude those that do not satisfy the necessary corner conditions, we detect only corners which are detectable under all reductions.
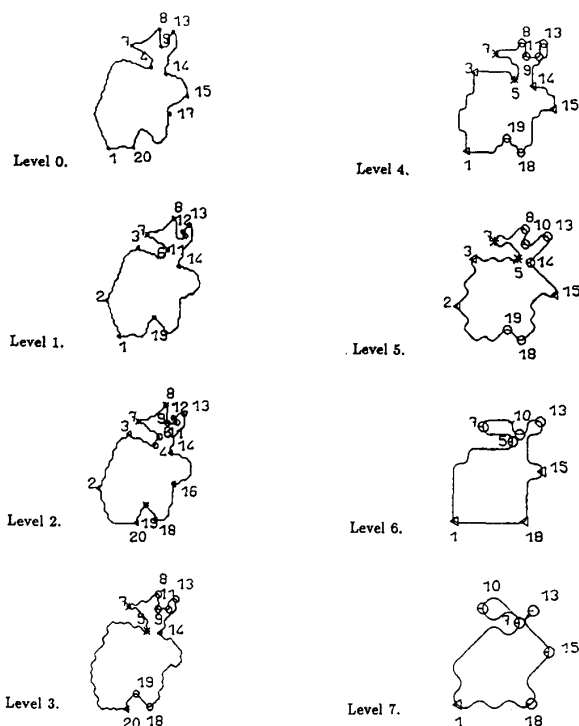
Fig. 5. Results of curve partitioning method by Fischler and Bolles.

Fig. 4. Corners of a closed curve in the pyramid: * indicates corners that satisfy the necessary corner conditions; ⊕ indicates corners which do not satisfy the distance condition; △ indicates corners not satisfying the condition of the angle's size, but satisfying the distance condition.

B. Corner Detection in the Pyramid

Since the description is a discrete one, the problems caused by undersampling have to be considered. It will be shown how this might affect the description, and a remedy will be presented.

In comparison to an analytical description of curves, where all curvature points that appear at a low resolution must also be detected at a high resolution, in pyramids curvature points may appear for the first time at a low resolution. These points correspond to curvatures which cannot be detected with the proposed parallel method because of the angle's size (for example, corner 15 or 18 in Fig. 4). Because of discrete sampling, corners may appear at level E_{i-1} and E_{i+1}, but not at level E_i (e.g., corner 3 is not detected at level 3 in Fig. 4). Furthermore, it is possible for a corner to be detected as two adjoining corners at the level above; but in these cases, the necessary corner conditions are not satisfied. However, the proposed algorithm does not suffer from such problems because information about corners is complemented with knowledge about the scale-space behavior of curves.

C. Robustness Measurements

Measures that reflect the size and scope of the curvature in the pyramid are defined next.

Three properties are important, namely, the lowest and highest levels at which a corner is detected, and at how many levels it is detected.

The last appearance (moving from the bottom to the top) of a corner in the pyramid gives information about the scope of the curvature; we therefore call it the *measure of scope* (S). The first appearance reflects the size of curvature; it is called the *measure of curvature-*

approximation (C). The sharper the enclosed angle, the earlier the curvature point will be detected. The number of levels at which a corner is detected is called the *measure of importance* (I). For individual corners, it was proved that they must be detected when they satisfy the corner conditions. The measure of importance describes the possibility of the appearance of a corner in the pyramid. It is introduced to make the description usable for more corners.

The three measurements stabilize the description in the sense that small changes in the input result in small changes in the description.

V. EXPERIMENTAL RESULTS

The results of applying our method to a curve (which was also used by Fischler et al. [6]) are shown in Fig. 4. The curve is plotted at eight successive levels of resolution. At every level, corners are extracted using the three-element method and classified in relation to size of angle and distance to the next corner.

The method of Fischler et al. for curve partitioning is based on the arc-chord distance. For every point on the curve, they decided whether the arc stays close to the chord or makes excursions away from it, and partitioned the curve at the points of single excursion that were farthest away from the chord. Points must be at least a predefined distance apart in order to be detected as different points. They chose one-quarter of a chord length as the threshold. In their paper, they showed the partitioning of the curve at one resolution; their results are shown in Fig. 5. Comparisons to our method show that with their relatively large threshold, they did not extract small corners (6, 11, 12, 17). They also did not detect point 2 and none of the points between 8 and 13.

The advantage of our method lies in the fact that not only corners at different resolutions are detected, but the descriptions at different resolutions are combined into one description. Therefore, corners can be differentiated by adding attributes to them.

VI. CONCLUSION

A multiresolution description of planar curves using corners and the curve pyramid has been presented. Continuous curves under smoothing have been examined, and the results used to define measures that stabilize the description. A method has been developed for detecting corners of digital curves in parallel. This local method has been analyzed; it was found that corners are detected in all cases when the straight lines enclose an angle of at least $63.4°$ $(108.4°)$ and the distance from one corner to the next is a receptive field (a receptive region of three cells).

A possible application of this description is multiple-resolution contour matching. Starting at a low resolution, the description can gradually be refined by adding the information stored at the next higher level of resolution of the pyramid.

REFERENCES

[1] H. Asada and M. Brady, "The curvature primal sketch," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 2–14, 1986.

[2] Z. Aviad, "A discrete scale-space representation," in *Proc. Int. Conf. Comput. Vis.*, 1987, pp. 417–422.

[3] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 26–33, 1986.

[4] F. Bergholm, "Edge focusing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 726–741, 1987.

[5] C. Fermüller, "Hierarchisches Vergleichen von Konturen," Master's thesis, Tech. Rep. 42, Inst. Image Processing Comput. Graph., Joanneum Res., Graz, 1989.

[6] M. A. Fischler and R. C. Bolles, "Perceptual organization and curve partitioning," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 100–105, 1986.

[7] M. A. Fischler and H. C. Wolf, "Saliency detection and partitioning planar curves," in *Proc. IEEE Image Understanding Workshop*, 1993, pp. 917–931.

[8] A. Goshtasby, "Multiple-scale segmentation and representation of solid plane shapes," in *Proc. Conf. Comput. Vis. Pattern Recogn.*, 1986, pp. 351–355.

[9] R. Hartley and A. Rosenfeld, "Hierarchical line linking for corner detection," Tech. Rep. CS-TR-1288, Center Automation Res., Univ. Maryland, 1983.

[10] J. J. Koenderik and A. van Doorn, "The structure of images," *Biolog. Cybern.*, vol. 50, pp. 363–370, 1984.

[11] W. G. Kropatsch, "A pyramid that grows by powers of 2," *Pattern Recogn. Lett.*, vol. 3, pp. 315–322, 1985.

[12] ——, "Curve representation in multiple resolution," *Pattern Recogn. Lett.*, vol. 6, no. 3, pp. 179–184, 1987.

[13] M. Leyton, "Symmetry-curvature duality," *Comput. Vis., Graph., Image Processing*, vol. 38, no. 3, pp. 327–341, 1987.

[14] T. Lindeberg, "Discrete scale-space theory and the scale-space primal sketch," Ph.D. dissertation, Computat. Vis. Active Percept. Lab., Roy. Inst. Technol., Stockholm, Sweden, 1991.

[15] P. Meer, E. S. Baugher, and A. Rosenfeld, "Extraction of trend lines and extrema from multiscale curves," *Pattern Recogn.*, vol. 21, no. 3, pp. 217–226, 1988.

[16] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 34–43, 1986.

[17] A. Rattarangski and R. T. Chin, "Scale-based detection of corners of planar curves," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 430–449, 1992.

[18] W. Richards and D. Hoffmann, "Codon constraints on closed 2D shapes," *Comput. Vis., Graph., Image Processing*, vol. 31, no. 3, pp. 265–281, 1985.

[19] W. Richards, B. Dawson, and D. Whittington, "Encoding contour shape by curvature extrema," *J. Opt. Soc. Amer.*, vol. 3, no. 3, pp. 1483–1491, 1986.

[20] T. Pavlidis, "Polygonal approximations by Newton's method," *IEEE Trans. Comput.*, vol. C-26, pp. 800–807, 1977.

[21] E. Saund, "Symbolic construction of a 2-D scale-space image," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 817–830, 1990.

[22] A. P. Witkin, "Scale-space filtering," in *Proc. 7th Int. Joint Conf. Artif. Intell.*, 1983, pp. 1019–1022.

[23] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 15–25, 1986.

Object Identification from Multiple Images Based on Point Matching Under a General Transformation

Mark C. K. Yang and Jong-Sen Lee

Abstract—This work is motivated by ship identification from a sequence of ISAR images. Maximum likelihood classification, based on point matching, is formulated when the observed images are subject to missing points and phantoms. The 3-D to 2-D transformation is assumed to be known only in a certain parametric form. Proper weights, based on the noise levels for all images, are derived for the classification formula. The new formulation simplifies the computation of matching and makes its extension to object identification from multiple images feasible. Moreover, some theoretical properties of the identification procedure can now be investigated. Guidelines on which groups of objects are easier to distinguish are found from statistical theory followed by intuitive explanation. This method is then applied to ship identification with simulated ISAR images.

Index Terms—Bayes rule, pattern recognition, point matching, multiple image information combination.

I. INTRODUCTION

Point matching plays an important role in pattern recognition. The work presented in this paper is to use salient points of objects to identify them from a sequence of pictures. The transformation from the 3-D points to a 2-D image is known only in a parametric form, and the unknown parameters may not be the same for different pictures. Moreover, the pictures are subject to noise so that some of the salient points may be missing and some false salient points, called phantoms, may be present in the pictures. If the viewing angle from the camera to the object is known, then we can consider this as a 2-D to 2-D point matching problem. Solution of this type of problem can be found in [1]–[3], where the basic idea is to examine all possible matchings by a fast algorithm which takes advantage of the relative distance between points. However, this method fails if the viewing angle between the camera and the object is unknown. Not only the profile of the object may change, but also the distance relation may no longer be preserved. Moreover, if the scaling is not the same along the x- and the y-axes, matching based on distance is more unworkable.

Flick and Jones [4] give a very general formulation of a d-dimension (d-D) to d–D mapping problem with the possibility of omission, phantom, and unknown transformation. Their basic idea is to write down the likelihood equation for the image based on a given object by all the possible matchings, and then to find the best parameters and the best matching using the maximum likelihood principle. However, due to the large number of parameters in the equation and the computational complexity in doing all the possible matchings, a solution is generally intractable unless the number of points is small. The single picture version of our problem can then be considered as a special case of the general setting in [4], but our formulation simplifies the computation and makes its extension to object identification from multiple images easier. Moreover, some theoretical properties of the identification procedure can now be investigated.