A Multiresolution Shape Description Algorithm

Gabriella Sanniti di Baja* and Edouard Thiel+

*Istituto di Cibernetica, CNR, 80072 Arco Felice, Naples, Italy
+Equipe TIMC-IMAG, CERMO BP 53 X, 38041 Grenoble cx, France

Abstract. A multiresolution shape description algorithm is presented, which is adequate to describe patterns perceived as the superposition of elongated regions. The weighted skeleton of the pattern is partitioned into a number of subsets, each corresponding to a pattern subset having simple shape, by means of a polygonal approximation. Different levels of description are possible, depending on the tolerance adopted during the approximation process. The computational cost of the algorithm is rather modest. A compact representation of the pattern is obtained, that takes simultaneously into account the representations of the pattern at the different levels.

1 Introduction

Shape description is an important step in a pattern recognition process. A multiresolution shape description scheme can be followed to facilitate recognition: a preliminary match can be attempted by comparing a rough description of the pattern at hand with the rough description of the prototypes; this allows one to reduce the number of prototypes in correspondence of which more fine details comparisons are necessary. Multiresolution can be obtained in two distinct ways: by varying the scale used by the adopted tool while the picture has fixed size, or by varying the resolution of the picture while the scale of the tool is fixed. In this paper we follow the first approach.

The structural approach is convenient to describe a pattern having complex shape: the pattern is decomposed into a number of simple regions that, by hypothesis, can be easily described by means of a suitable set of features; then, the description of the pattern is given in terms of the description of the obtained regions and of their spatial relationships. The same result can be obtained by decomposing, in place of the pattern, a suitable representation of it.

The decomposition should divide the representation system in such a way that each component could be interpreted as the representation of one of the simple regions into which the pattern is expected to be decomposed. Thus, the description of each region can be obtained by exploiting the information carried on by the corresponding component of the representation system. The choice of the representation system depends, besides the approach selected for shape description, also on the available computer architecture as well as on the shape structure of the pattern to be analysed.

A contour-based or a region-based system can be used to represent a single-valued pattern. A well-known region-based representation system, particularly suited to shape description [1-3], is the labelled skeleton. This is a thin subset of the pattern and its structure reflects the topological and geometrical features of the pattern. Each pixel of the skeleton is labelled with its distance from the complement of the pattern. Accordingly, it can be interpreted as the centre of a disc (whose shape depends on the adopted distance function) which fits the shape, the radius of the disc being proportional to the label of the pixel.

A correspondence can be established between any subset of the skeleton and the region of the pattern, which is the union of the discs associated with the pixels constituting the skeleton subset. This region can be recovered at a small computational cost by applying the reverse distance transformation to the skeleton subset. In fact, when conventional sequential computers are used, this process only requires two raster scan inspections of the array [4]. However, when regions associated to distinct skeleton subsets are of interest, the various skeleton subsets have to be individually subjected to the reverse distance transformation to avoid merging of the corresponding regions. The repeated application of the reverse distance transformation could excessively increase the overall computational cost, so preventing its use in the framework of a pattern decomposition process. Thus, it is desirable a skeleton analysis method which allows one to draw the regions associated with the skeleton subsets without resorting to the previous recovery process.

In this paper the skeleton of a pattern, perceived as constituted by the superposition of elongated regions, is interpreted as a curve in the 3D space, where the three co-ordinates of any pixel are the planar co-ordinates and the label. In the 3D space, the skeleton is partitioned by means of a polygonal approximation, in such a way that each partition component corresponds to a region with simple shape.

The obtained skeleton components are used to describe the corresponding regions. These can be drawn starting from the planar co-ordinates and labels of the extremes of the corresponding skeleton components, without resorting to the reverse distance transformation. Each region is constituted by the union of a central portion, which is trapezium-shaped, with two discs, whose diameters are equal to the bases of the trapezium. An example is shown in Fig.1. The centres of the discs are aligned along the skeleton component, which in turn is a symmetry axis for the trapezium. While the shape of the central portion does not depend on the distance function used to label the skeletal pixels, the discs are more circular if a quasi Euclidean distance is adopted.

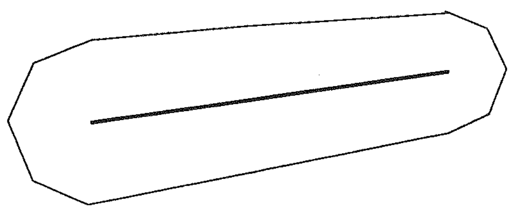

Fig. 1. A simple region associated with a skeleton component.

Reasonable approximations of the Euclidean distance are provided by the weighted distance functions [5, 6], where a suitable number of integer weights are used to measure the distance in between neighbouring pixels, depending on their relative positions. Here, we refer to a weighted distance involving two weights, respectively equal to 3 for the horizontal/vertical neighbours and 4 for the diagonal neighbours, and use the (3, 4)-weighted skeleton [3, 7] to represent the pattern at hand. Our description method has been inspired by a previous work [8], relative to the case of ribbon-like patterns, represented by the city-block distance labelled skeleton.

Pattern description is provided at different resolution levels, corresponding to different values for the threshold used during the polygonal approximation. The highest resolution level provides a fine details description and corresponds to the smallest threshold value. Each remaining level can be directly extracted from the highest level, without resorting to repeated application of the polygonal approximation. Thus, the computational cost to obtain the multiresolution description is practically the same as that necessary to get only the highest resolution description. The representations of the pattern at the various resolution levels are properly linked to each other, to provide a unique compact pattern representation.

2 The Weighted Skeleton

Let B and W be the two sets of black and white pixels constituting a binary picture digitised on the square grid. B and W are also referred to as the pattern and the complement. The 8-connectedness and the 4-connectedness are assumed for B and W, respectively. The 8-connectedness holds also for skeleton S, since this is a subset of B. We assume that a cleaning step aimed at removing the salt-and-pepper noise is accomplished on the picture, before computing the skeleton. In fact, loops originated in correspondence with non meaningful holes of B would irreparably bias the skeleton structure, and skeleton branches originating from non meaningful protrusions of B would require to include in the skeletonization algorithm also a pruning step. Cleaning will also prevent the existence of pathological patterns, as the lace-edged ones, that due to the complex structure of their contour could not be skeletonized.

The skeleton is a stick-like representation of the pattern which accounts for different shape properties, such as symmetry, elongation, width, and contour curvature. Research on skeletonization has been influenced, at least as concerns distance driven methods [7, 9-12], by the work of H. Blum involving the primitive notion of a symmetric point and a growth [13]. In a continuous pattern, a point p is called symmetric if at least two points of the boundary exist, such that their distance from p is equal to the distance of p from the boundary. For every symmetric point, the associated maximal disc is the largest disc, obtained as growth of the symmetric point, which is contained in the pattern. The set of symmetric points, each labelled with the radius of the associated maximal disc, constitutes the skeleton of the pattern. In turn, the pattern can be exactly reconstructed as the union of the maximal discs, the envelope of the discs being the pattern boundary.

The (3, 4)-weighted skeleton is the subset S of B having the following properties: 1) S has the same number of 8-connected components as B, and each component of S has the same number of 4-connected holes as the corresponding component of B. 2) S is centred within B. 3) S is union of simple 8-arcs and 8-curves. 4) The pixels of S are labelled with their (3, 4)-weighted distances from W. 5) S includes all the centres of the maximal discs of B, except for those removed to fulfil property 3).

We classify the pixels of S as end points, normal points and branch points, as follows. An end point is a pixel having a unique (4-connected) component of neighbours not in the skeleton. The end points identify the starting points of skeleton arcs and are placed in correspondence with the tips of pattern protrusions. A branch point is any pixel of S which is not an end point and has more than two neighbours in S. The branch points identify crossings of skeleton arcs, and are located in correspondence of superposition of regions of B. A normal point is a pixel of the skeleton which is neither an end point, nor a branch point. See Fig. 2.

A skeleton branch is an arc of the skeleton whose pixels are all normal points except for the two extremes. When all the pixels of the skeleton are normal points, the skeleton is a simple curve. In this event, to interpret the skeleton as constituted by branches, any pixel of the curve (e. g., the first skeletal pixel met by scanning the array in forward fashion) is taken to play the role of both the extremes of a looping skeleton branch.

We briefly outline the skeletonization algorithm [7] used in this paper to compute the (3, 4)-weighted skeleton. The process includes three steps: i) computation of the (3, 4)-weighted distance map DT, ii) identification of the set ML of the skeletal pixels, and iii) reduction of the set ML to the unit-wide skeleton S. The DT is a multi-valued replica of B, where every pixel is labelled with its (3, 4)-weighted distance from W. On the DT, the identification of the pixels of the ML (centres of maximal discs, saddle pixels and linking pixels) can be accomplished at a limited computational cost. In particular, one inspection of the DT is sufficient to identify all the centres of maximal discs and the saddle pixels. The identification of the centres of the maximal discs is a straightforward task since the label of any pixel is related to the radius of the associated disc. (The detailed method to correctly detect the centres of the maximal discs according to the (3, 4)-weighted distance can be found in [14]). The saddle pixels are identified by counting, in the neighbourhood of any pixel p, the number of components of neighbours with label respectively smaller and higher than p. The linking pixels necessary to guarantee the connectedness of the ML are found by growing paths along the direction of the steepest gradient in the DT, starting from any already found skeletal pixel. In the third step of the process, the set ML is reduced to unit width, by employing topology-and-end-point preserving removal operations.

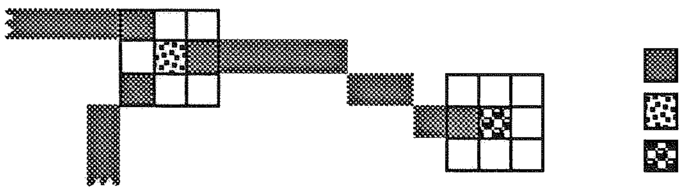

Fig. 2. Normal points, branch points and end points in the skeleton

Unless explicitly stated, in the following the distance label of any skeletal pixel is replaced by the normalised label. This can be obtained by dividing the label of the pixel by the weight used for the horizontal/vertical neighbours, in the adopted distance map. For the (3, 4)-weighted skeleton, the normalised label of a pixel with distance label p is the minimal integer k, such that $k \geq p/3$.

Using the (3, 4)-weighted distance is useful to obtain a skeleton stable under pattern rotation, which is an indispensable presupposition to produce the same decomposition whichever is pattern orientation. In fact, the discs provided by discrete metrics are polygons whose sides have fixed orientation. Thus the number of discs necessary to recover a pattern subset depends on geometry and orientation of the contour of the pattern subset. Skeleton stability increases with the number of sides characterising the disc. In this respect, discs built according to the (3, 4)-weighted distance have to be preferred to discs built according to the city-block or the chessboard distance. Stability regards, besides position and number, also the value of the distance labels associated with the skeletal pixels.

3 The Skeleton Partition

Aim of the partition process is to divide the skeleton into subsets such that each of them can be interpreted as the spine of a simple region. A region is simple if it satisfies the following two properties: 1) its local thickness is constant or changes monotonically and linearly along the spine; 2) the contour arcs of the subset which are common also to the contour of the elongated region are straight line segments. The spine of a simple region is a straight line segment along which labels are constant or monotonically and linearly change.

To identify the skeleton subsets representing simple regions, we perform a polygonal approximation on the skeleton. Indeed, if the skeleton is approximated in the (x, y) plane, only the property concerning with the rectilinearity of the contour of the represented regions is reflected by the obtained skeleton decomposition. To represent regions enjoying the property on thickness, as well as that on contour geometry, also the labels associated to the pixels of S have to be taken into account while performing the polygonal approximation.

Preliminarily, the skeleton is divided into its constituting branches. This is equivalent to perform a decomposition of the pattern into the elongated regions, that could be obtained by applying to the skeleton branches the reverse distance transformation.

Each skeleton branch is the spine of the corresponding elongated region. Then, on each branch a further partition is done to simulate the decomposition of the corresponding elongated region into subsets having simple shape. To this purpose, each branch is examined and interpreted as a curve in the 3D space. The planar co-ordinates and the normalised distance label of any skeletal pixel are the three co-ordinates (x, y, z) in the 3D skeleton representation. The branch is traced and a polygonal approximation is performed to partition the branch into a number of subsets which, in the limits of the adopted tolerance, are rectilinear 3D segments.

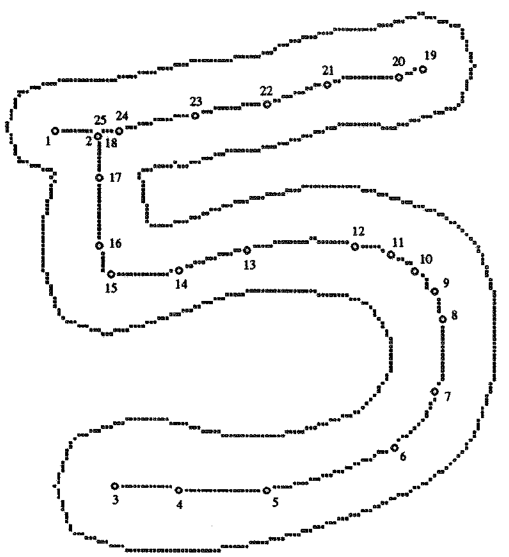

Fig. 3. Skeleton partition. The vertices, denoted by "o", are numbered in the order in which they are found during the polygonal approximation.

The approximation is accomplished by using a split type algorithm [15] in order the obtained set of vertices be not remarkably influenced by the sequenciality of the tracing process. The extremes of the current branch (say v_i and v_f) are accepted as vertices. Then, new vertices are identified in a recursive way. The Euclidean distance between any pixel of the skeleton branch and the 3D straight line (v_i, v_f) is computed, and the pixel v of the branch having the largest distance is identified. The pixel v is taken as a new vertex, if its distance from the straight line (v_i, v_f) is greater than a threshold, whose value depends on the adopted tolerance. Then, vertex selection is newly accomplished on the sub-branches v_i

v and v v_f. The recursive process terminates when the distance of all the pixels in the current sub-branch is below the threshold. As an example, see Fig. 3, where the skeleton partition has been obtained by using a threshold value equal to 1. 5.

The obtained components of the partition are used to extract geometric features of the represented regions. To this purpose, it is not necessary to recover the regions represented by the components of the partition by employing the reverse distance transformation. In fact, starting from the co-ordinates (x_1, y_1, z_1) and (x_2, y_2, z_2) of the vertices delimiting each partition component, an approximated version of the corresponding region can be built. This is obtained by drawing the octagon-shaped discs associated with the vertices, and a trapezium-shaped strip having the partition component as its symmetry axis.

In the past, algorithms to compute a multiresolution skeleton have been proposed [16-17]. There, by taking into account contour curvature to guide end point detection, different skeletons were available which differ from each other for the presence and/or for the length of some peripheral branches. Thus, a hierarchy among skeleton branches, and hence among elongated pattern subsets, was possible. Here, different representations of the same skeleton branch are obtained at different levels, by using different (increasing) threshold values to perform the polygonal approximation. The purpose is that of obtaining a multiresolution description of the elongated region associated with a skeleton branch. The lowest threshold produces the highest level (fine details) description.

Indeed, the polygonal approximation is performed only once, by employing the lowest value for the threshold. In fact, due to the use of a split type algorithm, candidates to be vertices at lower levels are pixels accepted as vertices at the highest level. The information necessary to identify the vertices of the successive approximations is available provided that as soon as a vertex is selected at a given step of the recursive process we record its distance. As the threshold increases, the description becomes rougher and rougher. In fact, whichever threshold is used, the region associated with a component of the skeleton partition is interpreted as consisting of the union of a trapezium-shaped strip with two discs. As an example, the regions obtained at four different resolution levels are shown in Fig. 4.

To use the various pattern descriptions by means of a compact representation, we associate each vertex of the polygonal approximation a quadruplet (x, y, label, permanence), where "permanence" accounts for the number of resolution levels at which the pixel is selected as a vertex. The compact pattern representation is illustrated in Table I, with reference to the 25 vertices found in the highest level decomposition of the skeleton shown in Fig. 3 and the successive four lower level pattern representations shown in Fig. 4. In Table I, the entries x, y, l, p, and t respectively indicate the Cartesian co-ordinates x and y, the normalised label, the permanence of the vertex in the successive approximations, and the pixel type (b, e, and n respectively stand for branch point, end point and normal point).

TABLE I

	x	y	l	p	t
1	30	40	15	5	b
2	20	38	12	5	e
3	35	128	16	5	e
4	51	128	17	1	n
5	73	129	14	4	n
6	105	118	12	5	n
7	116	104	14	3	n
8	117	86	14	5	n
9	115	79	16	1	n

	x	y	l	t	p
10	110	73	15	3	n
11	104	69	14	1	n
12	95	67	15	5	n
13	68	68	12	2	n
14	51	74	11	3	n
15	34	75	16	5	n
16	31	68	15	1	n
17	31	51	11	2	n
18	31	41	14	5	b

	x	y	l	p	t
19	112	22	14	5	e
20	106	25	13	1	n
21	88	26	13	1	n
22	73	32	12	2	n
23	55	34	14	1	n
24	36	39	13	1	n
25	31	40	15	5	b

The complexity of the skeleton decomposition algorithm is order of n. log(n), where n is the number of pixels of the skeleton. Using the normalised labels rather than the true distance labels allows us to apply the same polygonal approximation algorithm, whichever weighted distance function is used to label the skeletal pixels. Moreover, this enables us to treat uniformly both the planar and the label co-ordinates, since a displacement of at most one unit in each of the three directions is done when passing from a pixel of S to one of its neighbours along the skeleton branch. In this way, the 3D representation of the skeleton has the same topological properties as the 2D skeleton.

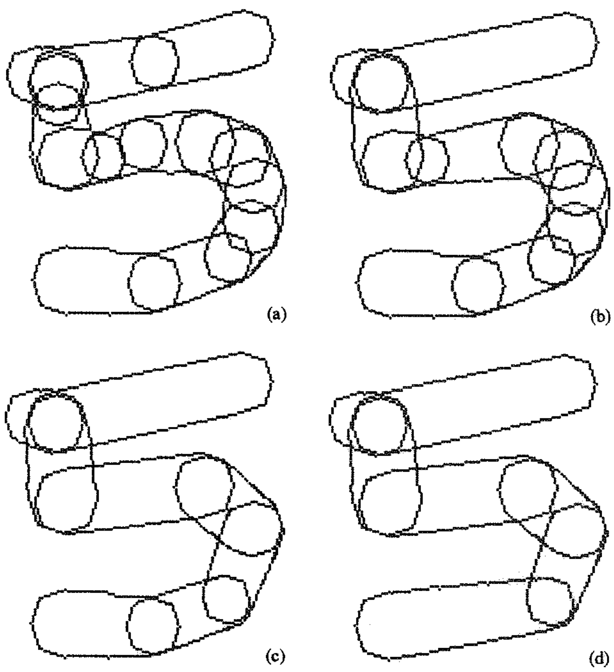

Fig. 4. Four decomposition levels corresponding to four approximations of the skeleton, respectively obtained by setting the threshold equal to 2. 5, (a), 4, (b), 6, (c), and 8, (d). The approximations are directly derived from the polygonal approximation done with threshold 1.5 (see Fig. 3).

4 Conclusion

A multiresolution description of the shape of a single-valued pattern is obtained starting from its weighted skeleton. The method is adequate for the description of patterns perceived as the superposition of elongated regions. The skeleton is partitioned into 3D rectilinear subsets by means of a polygonal approximation that takes into account both curvature and label variations. Since curvature and label variations reflect contour curvature variations and pattern thickness variations, the performed partition divides the skeleton into subsets corresponding to regions that, in the limits of the adopted tolerance, are characterised by linearly changing orientation and width. The value of the threshold used within the polygonal approximation conditions the number of partition components, as well as the quality of the obtained pattern decomposition. Thus, descriptions of the pattern at different resolution levels are possible. The highest resolution (fine details) level corresponds to the smallest threshold value. The vertices of the approximations at the lower levels can be directly identified while selecting the vertices at the highest resolution level. A compact representation of the pattern is obtained, that takes simultaneously into account the representations of the pattern at the different levels. This goal is reached by associating each vertex of the polygonal approximation corresponding to the highest resolution level a quadruplet (x, y, label, permanence) to record, besides the 3D co-ordinates, also the number of levels at which a pixel results to be a vertex of the approximating polygon.

The computational cost of the proposed method is order n. log(n). The use of a quasi Euclidean distance to label the skeletal pixels makes the decomposition of the skeleton (and hence of the pattern) stable under pattern rotation. Experimentation on the effects of rotation are currently under development.

References

1. H. Blum, R. N. Nagel: Shape description using weighted symmetric axis features. Pattern Recognition 10, 167-180 (1978)
2. A. Montanvert: Medial line - graph representation and shape description. Proc. 8th Int. Conf. on Pattern Recognition. Paris, 430-432 (1986)
3. E. Thiel, A. Montanvert: Shape splitting from medial lines using the 3-4 chamfer distance. In: C. Arcelli et al. (eds.): Visual Form Analysis and Recognition. New York: Plenum 1992, pp. 537-546
4. A. Rosenfeld, J. L. Pfaltz: Sequential operations in digital picture processing. J. ACM 13, 471-494 (1966)
5. G. Borgefors: Distance Transformation in Digital Images. Comput. Vision Graphics Image Process 34, 344-371 (1986)
6. E. Thiel, A. Montanvert: Chamfer masks - discrete distance functions, geometrical properties and optimization. Proc. 11th Int. Conf. on Pattern Recognition. The Hague, 244-247 (1992)
7. G. Sanniti di Baja: Well-shaped, stable and reversible skeletons from the (3, 4)-distance transform. Journal of Visual Communication and Image Representation. Accepted for publication (1992)
8. C. Arcelli, R. Colucci, G. Sanniti di Baja: On the description of digital strips. Proc. Int. Conf. on Artificial Intelligence Applications and Neural Networks. Zurich, 193-196, (1990)
9. L. Dorst: Pseudo-Euclidean skeletons. Proc. 8th Int. Conf. on Pattern Recognition. Paris, 286-288 (1986)
10. C. Arcelli, M. Frucci: Reversible skeletonization by (5, 7, 11)-erosion. In: C. Arcelli et al. (eds.): Visual Form Analysis and Recognition. New York: Plenum 1992, pp. 21-28
11. J. M. Chassery, A. Montanvert: Géométrie discrète en analyse d'images. Paris: Editions Hermès 1992
12. Y. Xia: Skeletonization via the realization of the fire front propagation and extinction in digital binary shapes. IEEE Trans. Patt. Anal. Mach. Intell 11, 1076-1086 (1989)
13. H. Blum: A transformation for extracting new descriptors of shape. In: W. Wathen-Dunn (ed.): Models for the Perception of Speech and Visual Form. Cambridge: M. I. T. Press 1967, pp. 362-380
14. C. Arcelli, G. Sanniti di Baja: Finding local maxima in a pseudo Euclidean distance transform. Comput. Vision Graphics Image Process. 43, 361-367 (1988)
15. T. Pavlidis: Structural pattern recognition. New York: Springer Verlag 1977
16. C. Arcelli: Pattern thinning by contour tracing. Comput. Graphics Image Process. 17, 130-144 (1981)
17. A. R Dill, M. D. Levine, P. B. Noble: Multiple resolution skeletons. IEEE Trans. Patt. Anal. Mach. Intell. 9, 495-504 (1987)