Pergamon

Pattern Recognition, Vol. 27, No. 8, pp. 1039–1049, 1994
Elsevier Science Ltd

0031-3203/94 $7.00 + .00

0031-3203(94)E0019-H

(3,4)-WEIGHTED SKELETON DECOMPOSITION FOR PATTERN REPRESENTATION AND DESCRIPTION

GABRIELLA SANNITI DI BAJA† and EDOUARD THIEL‡
† Istituto di Cibernetica, CNR, Via Toiano 6, 80072 Arco Felice, Naples, Italy
‡ Equipe TIMC-IMAG, CERMO BP 53, 38041 Grenoble Cedex 9, France

(*Received* 29 *May* 1993; *received for publication* 24 *February* 1994)

Abstract—A digital pattern, perceived as the superposition of elongated regions, is decomposed into simple regions through the decomposition of its (3,4)-weighted skeleton. The skeleton is interpreted as a curve in 3D space, where the three coordinates of any pixel are its planar coordinates and the distance label. The 3D curve is divided into rectilinear segments, which constitute the spines of elementary regions, i.e. regions with linearly changing width and orientation. Then, the spines are analysed to simplify the skeleton decomposition and avoid redundancy. Spines identifying regions unnecessary for the description of the pattern are annihilated, while contiguous spines, corresponding to sufficiently similar regions, are merged. The resulting skeleton components are used to represent and describe the simple regions into which the pattern is decomposed. Decomposition at different resolution levels can be obtained by selecting different threshold values during the polygonal approximation, performed to divide the skeleton into rectilinear pieces, and/or the successive merging step.

Weighted distance Labelled skeleton Polygonal approximation Decomposition Description

1. INTRODUCTION

The description of patterns that can be perceived as the superposition of ribbon-like regions[1] can be facilitated by the structural approach. A suitable representation of the pattern is decomposed, in such a way that each decomposition component could be interpreted as the representation of one of the regions constituting the pattern.[2–7] Then, the description of the pattern is obtained in terms of the description of the obtained regions and of their spatial relationships. The description of each elementary region, which by hypothesis is characterised by a simple shape, can be obtained by exploiting the information carried on by the corresponding component of the representation system.

The labelled skeleton[8–13] is a convenient tool to analyse the shape of patterns perceived as union of ribbon-like regions. A ribbon-like pattern is characterised by one spine and a disc: the disc sweeps out the shape by moving along the spine, changing size as it moves. The skeleton is a curvilinear subset of the pattern and its branches play the role of the spines of the ribbon-like regions constituting the pattern. The label of any pixel p of the skeleton, which represents the distance of p from the complement of the pattern, can be interpreted as the radius of the sweeping disc centred on p. The shape of the disc depends on the adopted distance function. The discs are more rounded if a quasi Euclidean metric is adopted.

Reasonable approximations of the Euclidean distance are provided by the weighted distance functions,[14–17] where suitable integer weights are used to measure the distance between neighbouring pixels, depending on their relative position. Skeletons[7,10,12,13] whose pixels are labelled using a weighted distance are called weighted skeletons. The weighted skeleton and the skeleton labelled according to the city-block or the chessboard distance can be obtained at a comparable (limited) computational cost. The stability of the weighted skeletons under pattern rotation favours their use for practical applications.

A correspondence exists between any subset of the skeleton and the region of the pattern that is the union of the discs associated with the pixels of the skeleton subset (in the strict sense, the only pixels of the skeleton subset which are centres of maximal discs are enough to recover the region). This region can be obtained by applying the reverse distance transformation[18,19] to the skeleton subset, which requires two raster scan inspections when a sequential algorithm is used. Under certain circumstances, a satisfactory approximated version of the region can be obtained at a lower computational cost. For instance, if the skeleton subset can be interpreted as the spine of an elementary region having linearly (and monotonically) changing width and orientation, a satisfactory approximated region is the envelope of only two discs, those associated with the extremes of the spine.

In this paper, we divide the (3,4)-weighted skeleton[7,13] of a pattern into subsets that can be understood as spines of simple regions. The decomposition method has been inspired by previous works[5,6] where the city-block distance labelled skeleton has been employed. Skeleton decomposition is accomplished in two main

phases. During the first phase (splitting step), the skeleton is interpreted as a curve in 3D space, where the coordinates of any pixel are the planar coordinates and the distance label. The 3D skeleton is divided into rectilinear segments by means of a polygonal approximation. The obtained segments are the spines of elementary regions, characterised by linearly (and monotonically) changing width and orientation. During the second phase (merging step), the skeleton segments are analysed. Skeleton segments which represent regions almost completely recovered by adjacent regions are annihilated, while contiguous segments constituting the spines of sufficiently similar elementary regions are merged. The merging step reduces the number of regions into which the pattern is decomposed, so that the obtained results are more in accordance with human intuition. In particular, it allows alleviation of the distortions generally affecting geometry and labels of the skeleton in correspondence of region crossings, which could produce an error-prone decomposition.

Although referred to the case of patterns represented by their (3,4)-weighted skeleton, the decomposition process equally applies to any other weighted skeleton. In fact, most of the computation is done by using normalized labels, which rids the process of the dependence on the value of the weights of the adopted distance function.

2. NOTIONS AND DEFINITIONS

Let B and W be a pattern perceived as the superposition of elongated regions, and its complement, respectively. Since the tool we adopt to perform the decomposition of B is the weighted skeleton S, we assume that a cleaning step is preliminarily performed to fill the noisy holes of B. In fact, loops originated in S in correspondence with non-meaningful holes of B would irreparably bias the skeleton structure and strongly condition the resulting decomposition. Moreover, the existence of noisy holes, completely lacing the contour of B, could even prevent creation of a unit-wide skeleton. We do not explicitly require that cleaning also remove noisy protrusions and dents from the contour of B, so as to avoid creation of noisy skeleton branches. These branches can be removed after the skeleton has been obtained, during the pruning step that we performed to delete skeleton branches regarded as non-significant in the problem domain.

The 8-connectedness and the 4-connectedness are assumed for B and W, respectively. The 8-connectedness holds also for S, since the skeleton is a subset of B.

The (3,4)-weighted distance $d_{3,4}$ among two pixels p and q is the length of the shortest 8-connected path (not necessarily unique) from p to q, where the two integers weights $w_1 = 3$ and $w_2 = 4$ are used to measure any horizontal/vertical unit move and any diagonal unit move, respectively.

The distance transform DT of B with respect to W is a replica of B, where each pixel is labelled with its $d_{3,4}$ distance from W. Each pixel of the DT can be interpreted as the centre of a disc, whose radius has length equal to the label of the pixel. The disc is octagon-shaped, as soon as the centre has a sufficiently large label.

A disc of the DT which is not completely overlapped by any other single disc is called a maximal disc. The union of the maximal discs coincides in size and shape with B. The centre of a maximal disc is called a maximal centre. A suitable comparison[20] among the label of a pixel p and the label of its neighbours in the DT allows one to establish whether p is a maximal centre.

The (3,4)-weighted skeleton S is the subset 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 the unit-wide union of simple 8-arcs and 8-curves.[21] (4) The pixels of S are labelled with their $d_{3,4}$ distance from W. (5) S includes almost all the maximal centres of B [complete inclusion is not compatible with fulfilment of property (3)].

An end point is a pixel of S having a unique (4-connected) component of neighbours not in the skeleton. It identifies the starting point of a skeleton arc and is placed in correspondence with the tip of an elongated subset of the pattern. A branch point is a pixel of S which is not an end point and has more than two neighbours in S. It identifies crossings of skeleton arcs, and is located in correspondence of the superposition of elongated regions. Finally, a normal point is a pixel of the skeleton which is neither an end point, nor a branch point (see Fig. 1).

Branch points and end points allow one to interpret the skeleton as a concatenation of skeleton branches. These are arcs of the skeleton, whose pixels are all normal points except for the extremes. In particular, branches delimited by an end point are called peripheral skeleton branches. When all the pixels of S are normal points, the skeleton is a simple curve. The curve is interpreted as a single skeleton branch, whose extremes are any two adjacent normal points, e.g. the first pixel of S, met when scanning the array in forward raster fashion, and the last pixel found when tracing S completely, starting from the first extreme. These two pixels are considered as if they were branch points, so as to treat all the skeleton branches homogeneously.

The normalized label of a skeleton pixel with distance label p is the minimal integer k, such that $k \geq p/3$. Unless differently specified, in the following the same letter will be used to indicate both the pixel and its associated label.

An elementary region of B is any set R, obtained by applying the reverse distance transformation to a (subset of a) skeleton branch, such that: (1) the local thickness of R changes monotonically and linearly along the (subset of the) skeleton branch; and (2) the subsets of the contour shared by R and B are straight line segments.

The (subset of a) skeleton branch corresponding to the elementary region R is the spine of R.

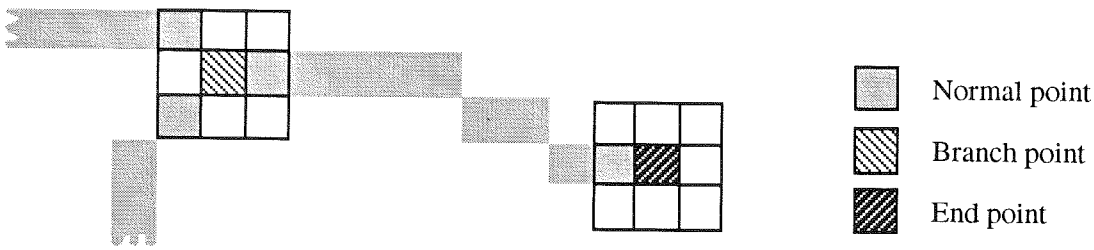

Fig. 1. Normal points, branch points and end points in the skeleton.

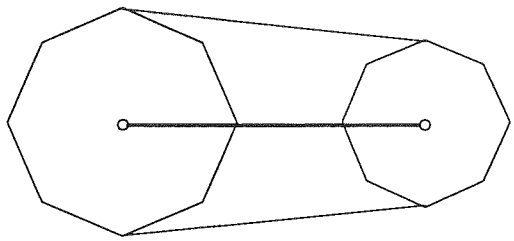

Fig. 2. The approximated version of an elementary region, obtained by building the envelope of the two octagon-shaped discs, centred on the extremes of the corresponding skeleton subset.

A satisfactory approximated version of R can be obtained by taking the envelope of the discs associated with the extremes of the spine of R. The two discs are octagon-shaped, while the central portion of the envelope is trapezium-shaped (see Fig. 2).

3. DECOMPOSING THE SKELETON: THE SPLITTING STEP

The skeleton is preliminarily processed to prune skeleton branches originated by protrusions regarded as non-significant in the problem domain. These branches include, but are not limited to, branches originated from noisy contour protrusions. In fact, pruning is also effective to reduce skeleton sensitivity to pattern rotation, by removing skeleton branches whose presence depends on pattern orientation.

Pruning is sometimes already included as a final step of the skeletonisation algorithm. This is the case in the algorithm proposed by Sanniti di Baja,[13] where pruning is accomplished using a criterion based on the relevance of the protrusion associated with the skeleton branch, so as to keep under control the loss of information caused by branch deletion.

In this paper, the above relevance criterion is used, while pruning is parallelwise applied to the peripheral

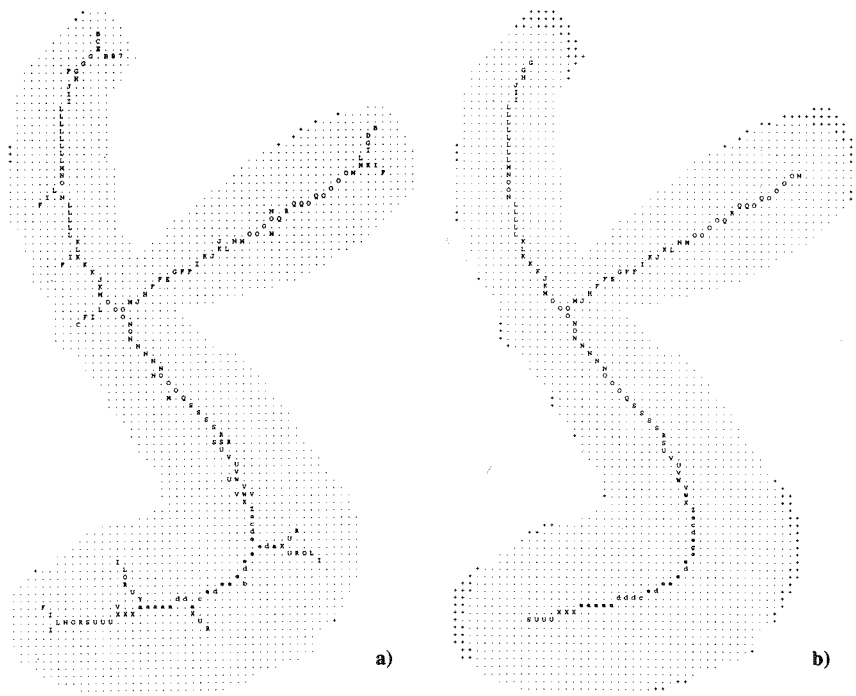

Fig. 3. The (3,4)-weighted skeleton of a pattern before pruning (a) and after pruning (b). The pixels of the input pattern which are not recovered by applying to the skeleton the reverse distance transformation are indicated by "+".

skeleton branches. Pruning is iterated until branch removal does not diminish the representative power of the skeleton, in the limits of the adopted tolerance. Thus, skeleton branches which are initially delimited by branch points may be pruned as soon as they become peripheral branches, due to the deletion of the neighbouring branches. Iterating pruning does not cause a summation effect in the loss of information. At each iteration and for each peripheral skeleton branch, the protrusion whose relevance is evaluated is the protrusion mapped in the union of the current peripheral skeleton branch with the neighbouring skeleton branches, already pruned at a previous iteration. To this purpose, the information relative to the starting point(s) of the branch(es) is propagated through the branch(es) while performing pruning.

In Fig. 3, the (3,4)-weighted skeleton is shown superimposed on the input pattern, before and after pruning. Pixels denoted by "+" are not recovered when the reverse distance transformation is applied to the skeleton. In Fig. 3(a), only a few pixels of the border of B are missed out. Loss of recovery happens since the skeleton is required to be unit-wide and, as such, it does not include all the maximal centres of B. In Fig. 3(b), the smoothing effect due to the pruning process is evident; each skeleton branch remaining after pruning corresponds to a significant ribbon-like region.

The skeleton is preliminarily decomposed into the constituent skeleton branches. This is equivalent to performing a decomposition of the pattern into the elongated regions that could be obtained by individually applying the reverse distance transformation to the skeleton branches. A data structure is built to record the extremes of the skeleton branches and the spatial relationships among them.

Each skeleton branch is furthermore decomposed, by means of a polygonal approximation, in such a way that each rectilinear segment constitutes the spine of an elementary region. Division points have to be placed wherever non-linear curvature changes occur along the skeleton branch, as they reflect non-linear curvature changes along the contour of the corresponding pattern subset. Division points have also to be placed where non-linear or non-monotonic label variations occur, as they indicate non-linear or non-monotonic pattern thickness variations. To locate both types of division points, we interpret any skeleton branch as an arc in 3D space where, for each skeletal pixel, the three coordinates are the planar coordinates and the normalised label. Using the normalised label in place of the distance label is done to treat uniformly the three coordinates, by allowing a displacement of one unit only in each of the three directions, when passing from a skeletal pixel to one of its neighbours. In this way the skeleton branch is a connected arc also in the 3D representation.

The polygonal approximation is accomplished by using a split type algorithm (e.g. the one described by Pavlidis,[22]), so that the obtained set of vertices is not influenced by the order in which skeletal pixels are processed. The extremes of the current branch (say v_i and v_f) are accepted as vertices, and properly stored in the data structure. Then, new vertices are identified (and stored in the data structure) in a recursive way. The Euclidean distance $d_E(p)$ between any pixel p of the skeleton branch and the 3D straight line (v_i, v_f) is computed. Then, the pixel of the branch for which $d_E(p)$ has the largest value is taken as a new vertex v, provided that $d_E(p)$ is greater than an a priori fixed threshold θ. Vertex selection is then accomplished on the sub-arcs $v_i v$ and $v v_f$. The recursive process terminates when, for the pixel maximising $d_E(p)$, it results in $d_E(p) \leq \theta$.

If a number of pixels of the skeleton branch maximises $d_E(p)$, the two pixels which are respectively the closest to v_i and to v_f are accepted as vertices, and the skeleton arc is divided in three sub-arcs, which are recursively examined. Accepting as vertices all the pixels maximising $d_E(p)$ could result in a polygonal approximation with too many vertices, not necessarily all significant. In contrast, accepting only one pixel could make the skeleton decomposition dependent on the order in which the pixels are processed.

The square root computation necessary to obtain the Euclidean distance $d_E(p)$ can be avoided, since the same result is obtained when comparing the square distance with the square threshold.

The value of θ is fixed depending on the tolerance regarded as acceptable for the specific task. The threshold should be rather small, to favour a quite faithful recovery of the elementary regions having the skeleton segments as their spines, by building the envelopes of the pairs of discs, centred on the extremes of the spines. In our experiments, the value $\theta = 1.5$ has been revealed as adequate. Skeleton decomposition at different resolution levels is obtained by assigning different values to θ. As the threshold increases, the number of components into which the skeleton is decomposed generally diminishes, while the representation becomes rougher and rougher. In fact, the regions that could be recovered by applying the reverse distance transformation to the pixels of the skeleton components are likely to differ remarkably from the envelopes of the discs centred on the extremes of the so-found skeleton components.

Let $V(\theta)$ be the set of vertices found in the polygonal approximation of the skeleton, performed with the lowest threshold. The vertices of any other polygonal approximation, performed with a higher threshold, can be directly identified as they constitute a subset of $V(\theta)$. Any subset is obtained by comparing the value $d_E(p)$, stored for any pixel p of $V(\theta)$, with the desired new threshold.

Recovery of the elementary regions represented by the skeleton segments is not necessary for computing geometric features (e.g. area or perimeter) and shape features (e.g. orientation or rectangularity) of the regions. These features can easily be derived starting from the 3D coordinates of the found vertices. However, for illustrative purpose, the elementary regions, corresponding to polygonal approximations of the skeleton at three different threshold values, are shown in Fig. 4.

Fig. 4. Letter V denotes the vertices found during the polygonal approximation of the skeleton with threshold $\theta = 1.5$ (a); decomposition into elementary regions, corresponding to $\theta = 1.5$ (b), $\theta = 4$ (c), and $\theta = 8$ (d).

To use all the pattern representations in a compact way, we associate each vertex p of $V(\theta)$ a quadruplet $(x, y, \text{label}, d_E(p))$. In this way, the permanence of a vertex in any of the resolution levels can be immediately checked. In Table 1, the entries x, y, l, d, and t indicate the Cartesian coordinates x and y, the label, the 3D distance, and the pixel type, respectively (b, e, and n, stand for branch point, end point and normal point,

Table 1.

	x	y	l	d	t
1	31	22	16		e
2	28	28	21	1,99	n
3	28	40	23	5,23	n
4	30	48	20	1,86	n
5	35	55	24		b
6	31	97	28		e
7	46	94	38	2,94	n
8	53	88	40	21,24	n
9	49	72	27	4,60	n
10	36	56	24		b
11	67	37	23		e
12	57	43	26	2,24	n
13	41	51	15	3,15	n
14	36	55	24		b

respectively) with reference to the 14 vertices, found when performing the polygonal approximation with $\theta = 1.5$ of the skeleton shown in Fig. 4(a).

4. DECOMPOSING THE SKELETON: THE MERGING STEP

Some of the regions represented by the skeleton segments are (almost completely) overlapped by the adjacent regions, especially when the polygonal approximation is performed with a low threshold value. To avoid redundancy, overlapped regions should be either completely disregarded or merged to adjacent regions. To this purpose, the skeleton segments must be examined and suitably processed. Correspondingly, the data structure, where the vertices found on the skeleton branches have been recorded, is suitably updated.

4.1. *Short spine annihilation*

All the short spines (e.g. skeleton segments having length less than 4 pixels) are examined. Short spines, whose contiguous spines are not short, are termed isolated short spines.

Any isolated short spine s_i, delimited by two normal points, is annihilated by moving the two vertices, shared by s_i with the contiguous spines s_{i-1} and s_{i+1}, towards a common position in 3D space. This position is either the barycentre of s_i or the intersection between s_{i-1} and s_{i+1}, depending on the angle between s_{i-1} and s_{i+1}. The pixel common to the two new spines, obtained by modifying s_{i-1} and s_{i+1}, is not taken as a vertex if, in the limits of the tolerance adopted when performing the polygonal approximation, the modified spines are aligned. When this is the case, the two modified spines are merged into a unique spine, still identifying an elementary region.

Any isolated short spine s_i, delimited by end points and/or branch points, is suppressed. However, when s_i is delimited by a branch point, track of s_i must be kept to record the relative position of the skeleton branches and, hence, of the corresponding regions in the decomposed pattern. In this case, the suppressed spine s_i plays the role of a linking element, but has no region representation power in the decomposition.

4.2. *Superfluous spine annihilation*

Spines longer than 4 pixels may correspond to regions almost completely overlapped by adjacent regions. These spines are superfluous for pattern representation and description.

In general, a spine can be considered superfluous if the envelope of the discs centred on its extremes does not significantly differ from the union of the two discs, as it is the case if the two discs partially overlap. The $d_{3,4}$ distance between the extremes of the spine and the value of their labels, e.g. l_1 and l_2, can be used to evaluate the overlapping. The spine is regarded as superfluous if the overlapping condition is satisfied, i.e. if it results in $l_1^2 + l_2^2 \geq d_{3,4}^2$.

If the extremes of a superfluous spine s are both normal points and the contiguous spines are both non-superfluous, s is annihilated by following the same strategy already discussed for the short spine annihilation.

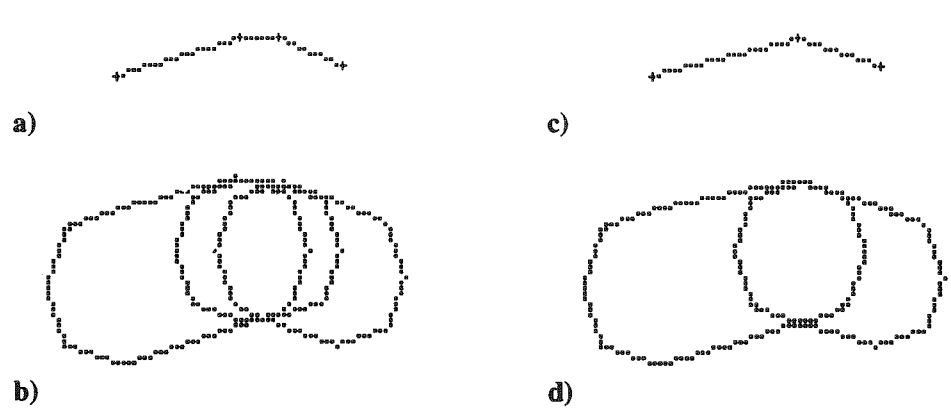

Fig. 5. The initial partition of the skeleton into three components (a), originates a decomposition (b), where the intermediate region is not significant. After the intermediate spine is annihilated (c), a more significant decomposition is obtained (d). The vertices of the partition components are denoted by "+".

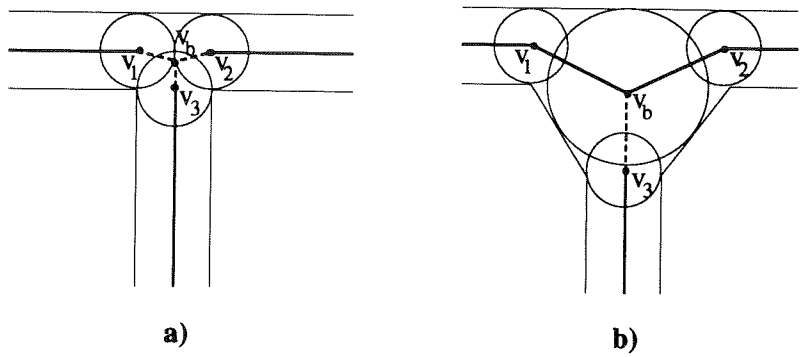

Fig. 6. Skeleton segments rid of region representation power are hatched lines. The overlapping condition is satisfied by the discs centred on the pairs of vertices (v_1, v_3) and (v_2, v_3) in (a), and by the discs centred on the pair (v_b, v_3) in (b).

The effect of the spine annihilation process is illustrated in Fig. 5.

4.3. *Crossing spines*

Spines sharing a branch point as a common vertex identify elementary regions that cross each other. The spines are meaningful if the corresponding elementary regions do not remarkably overlap. Otherwise, their presence in the skeleton is necessary only to keep track of the spatial relationships among skeleton branches. As before, these spines should simply play the role of linking elements, but have no region representation power in the decomposition.

However, it may happen that considering any such a spine as a linking element is not enough to eliminate redundancy in the decomposition, and to guarantee stability of the decomposition under pattern rotation. Also, some successive spine(s), located on the same skeleton branch, might have no region representation power. The skeleton branch could have been excessively fragmented during the polygonal approximation, so that the subset of the skeleton branch which, in the final decomposition, should have only role of a linking element results in being divided into a number of spines. Moreover, the polygonal approximation of the skeleton performed at the lowest threshold is seldom stable when the orientation of the input pattern is changed. Thus, a different interpretation of which spines are linking elements could occur when the pattern is rotated.

Pixels of $V(\theta)$ which are vertices of the polygonal approximation also when a larger threshold is selected ($\theta = 2$, in our case) are more likely both to be present in the skeleton decomposition when the pattern is rotated, and to identify correctly the subset of the skeleton branch having only a linking role in the final decomposition. These pixels are identified by resorting to the quadruplets $(x, y, \text{label}, d_E(p))$, stored for any vertex found during the polygonal approximation.

Let $b_1, b_2, \ldots, b_n$ be the skeleton branches sharing the branch point v_b, and let v_i $(i = 1, n)$ be the first vertex along b_i, remaining in the polygonal approximation when it is $\theta = 2$. The subset of b_i, delimited by v_b and v_i, is regarded as rid of region representation power if for a vertex v_j, located on b_j $(j \neq i)$, the overlapping condition between the discs centred on v_j and v_i is satisfied. If this is the case, the subset only maintains its linking role. Otherwise, the overlapping condition is checked between the discs centred on v_b and on any of the vertices of the subset of b_i, delimited by v_b and v_i, to identify the longest linking element and to simplify the decomposition.

As an example, refer to Fig. 6.

4.4. *Spine merging*

Although the spines remaining at this stage of the process all significantly contribute to pattern recovery, merging some of them could be useful to reduce the number of regions which will constitute the primitives for pattern decomposition. The reduction of the number of primitives generally also increases the stability of the decomposition and produces results more in accordance with human intuition.

Elementary regions having sufficiently similar width and orientation could be merged by merging the corresponding spines. A merged region, although no longer an elementary region, can still be simply described starting from the 3D coordinates of the vertices of the merged spines. By employing a different merging tolerance, different concatenations of merged spines are possible, which produce different pattern decompositions. The decompositions have all the same representation power, since any merged region is the union of the corresponding elementary regions. (This was not the case when different decompositions were obtained by adopting different thresholds during the polygonal approximation of the skeleton, see Fig. 4.)

Having several decompositions of the same pattern is convenient to facilitate pattern recognition, and allows one to select the decomposition which is more adequate to the solution of a specific problem.

The orientation and width of two contiguous elementary regions are faithfully reflected by their corresponding spines. Thus, if the spines, represented as 3D straight lines, are aligned in the limits of the adopted tolerance, the elementary regions are similar and could be merged.

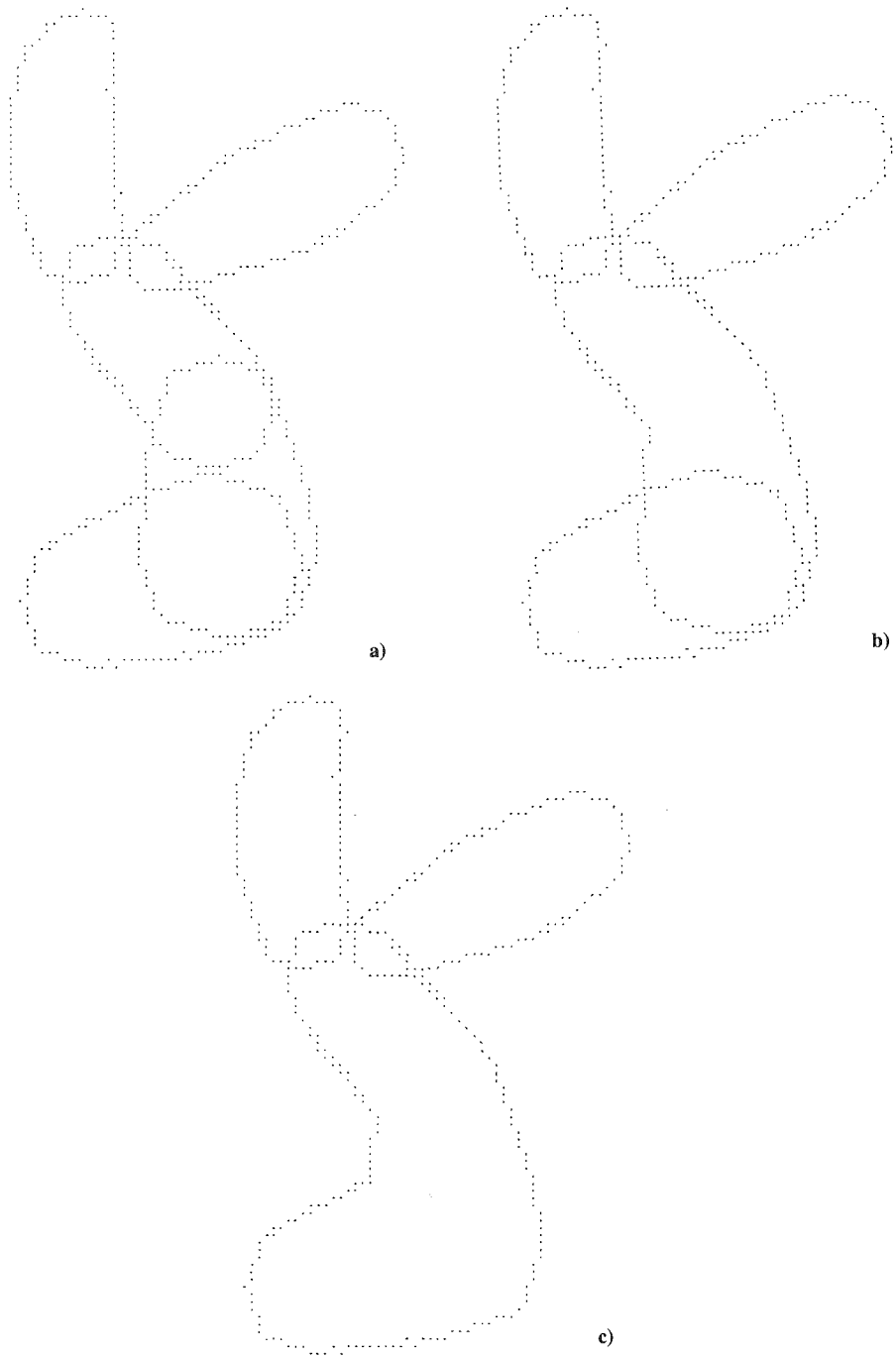

Fig. 7. Different decompositions of the same pattern, obtained by using different values for the merging threshold: $\tau = 0.15$ (a), $\tau = 0.25$ (b), $\tau = 0.50$ (c).

Each pair of successive spines, belonging to the same skeleton branch, is examined. Let (v_{i-1}, v_i) and (v_i, v_{i+1}) be the vertices delimiting the current pair. Let D_i and L_i be the Euclidean distance of v_i from the straight line segment joining v_{i-1} and v_{i+1}, and the Euclidean length of the segment, respectively. A flag F, initially equal to 0, is set to 1 in correspondence with each vertex v_i, such that D_i/L_i is less than an a priori fixed merging threshold τ.

Let $v_1, v_2, \ldots, v_n$ be a set of successive vertices, in correspondence of which it is $F = 1$. Moreover, let v_0 and v_{n+1} be the vertices immediately preceding v_1 and immediately following v_n.

If $n = 1$, the two spines (v_0, v_1) and (v_1, v_{n+1}) are merged by all means.

If $n > 1$, the distance D_i from the straight line segment joining v_0 with v_{n+1} is divided by the length L_i of the segment, for every v_i $(i = 1, 2, \ldots, n)$. If $D_i/L_i < \tau$ for

every vertex, all the spines are merged. Otherwise, the concatenation $v_2, v_3, \ldots, v_{n-1}$ is considered, and for each of these vertices the merging ratio D_i/L_i is checked with reference to the straight line segment joining v_1 and v_n. The process is repeated until for the concatenation $v_k, v_{k+1}, \ldots, v_j$ $(k = 1 + i, j = n - i, i > 0)$ the merging condition is verified by all the vertices. Then, the merging condition is recursively checked on the two sub-concatenations $v_1, v_2, \ldots, v_{k-1}$ and $v_{j+1}, v_{j+2}, \ldots, v_n$.

The vertices delimiting the set of the merged successive spines are taken as the extremes of the resulting complex spine. Note that the remaining vertices still maintain their region representation power, since the region associated with a complex spine is the union of the elementary regions associated with the merged spines.

The value of the merging threshold τ depends on the desired merging tolerance. In our experiments, the value $\tau = 0.25$ has been adopted as a default value. Larger values can be used to favour merging. An example is shown in Fig. 7, where three different values have been used for the merging threshold τ. The three decompositions are obtained starting from the polygonal approximation of the skeleton, performed with $\theta = 1.5$. Note that, in contrast to the decompositions shown in Fig. 4, the regions are not elementary regions.

The possibility of merging spines sharing a branch point as a common vertex could also be taken into account, so that the final pattern decomposition would not be conditioned by the preliminary decomposition of the skeleton into its constituting branches. Work in this respect is currently in progress.

5. CONCLUSION

In this paper we have illustrated a method for decomposing a digital pattern through the decomposition of its weighted skeleton. The method is adequate for patterns that can be perceived as constituted by the union of elongated (ribbon-like) regions; it could be employed, for instance, in the framework of a document analysis task to classify the alphanumeric symbols which it contains.

The weighted skeleton has been chosen to favour the stability of the decomposition under pattern rotation. In fact, stability is an indispensable presupposition for

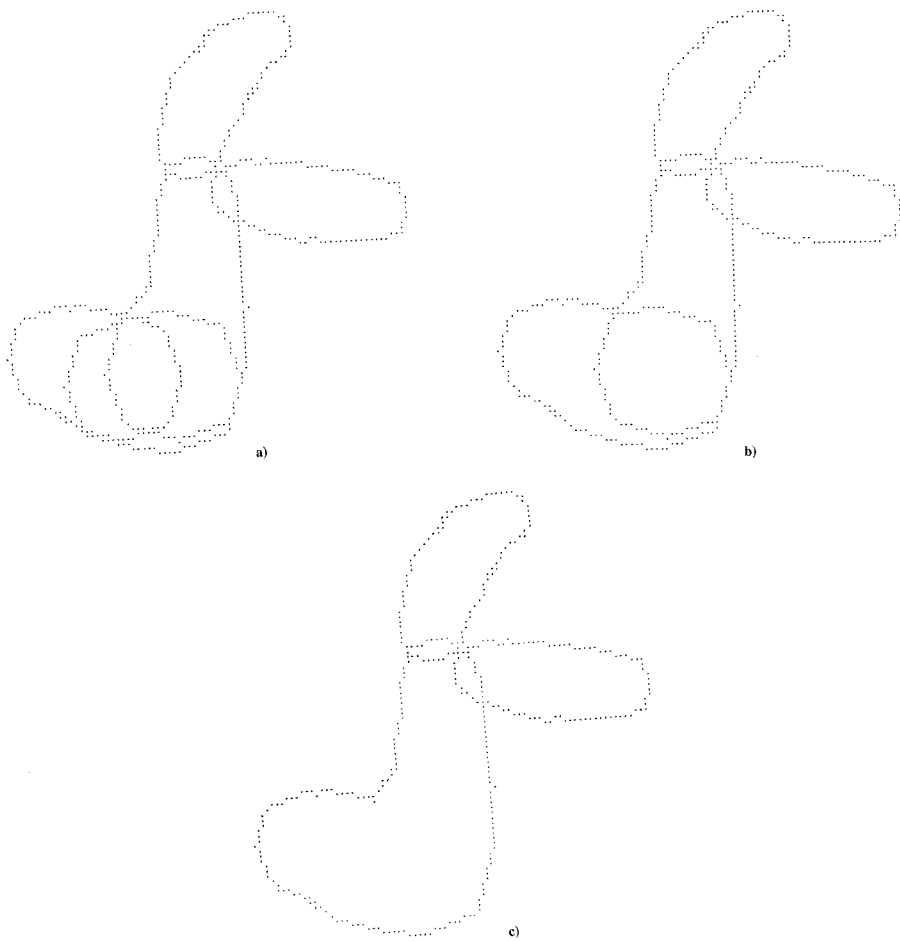

Fig. 8. Stability of the decomposition under pattern rotation.

any application where the orientation of the pattern is not known a priori. Stability is also favoured by the annihilation and merging steps, which reduce the skeleton decomposition components to the most significant ones. As an example of the performance of the decomposition under pattern rotation refer to Fig. 8, where the pattern of Fig. 7 appears in a different orientation. The three decompositions have been obtained using the same merging threshold as in Fig. 7.

A relevant feature of the proposed decomposition method is the possibility of obtaining decompositions at different resolution levels. This can be done by changing the thresholds used during the polygonal approximation and the merging step. In the first case, the obtained pattern representations do not have the same representative power. In fact, the skeleton decomposition components are considered as the spines of elementary regions, independently of the employed threshold. The various representations can be used in a compact way. In the second case, the decompositions differ from each other for the number and shape of the constituent regions, but all have the same representative power. A region of a decomposition obtained with a small merging threshold is simpler to describe, but the total description of the pattern in terms of the constituent regions is less manageable.

The computational burden of the process is rather modest, because all the computations are performed on a small amount of data (the skeletal pixels and, afterwards, the vertices of the polygonal approximation), which are stored in vector form.

We are conscious that our method can be improved, especially regarding the merging step. It should be interpreted as a starting point to devise better decomposition procedures, each tailored to the specific application. It is, in fact, rather difficult to foresee a general-purpose decomposition process. For completeness, we point out some of the topics that we are currently investigating. We are trying to take into account some more information, still derivable from the coordinates of the vertices of the polygonal approximation of the skeleton, which could originate merged regions more in accordance with human intuition. This information concerns detection of vertices where the change in sign of the curvature occurs, or the distance label is minimal. Detecting these pixels could help us to avoid considering S-shaped patterns or clepsydra-shaped patterns as a unique region. Also, the splitting step could be performed by resorting to a skeleton partition more appropriate to handling rounded shapes. For instance, rather than a polygonal approximation, one could employ a curve-fitting technique.

REFERENCES

1. A. Rosenfeld, Axial representation of shape, *Computer Vision Graphics and Image Processing* **33**, 156–173 (1986).
2. H. Blum and R. N. Nagel, Shape description using weighted symmetric axis features, *Pattern Recognition* **10**, 167–180 (1978).
3. L. P. Cordella and G. Sanniti di Baja, An approach to the decomposition of complex figures, *Digital Image Analysis*, S. Levialdi, ed., pp. 155–162. Pitman, Tunbridge wells (1984).
4. A. Montanvert, Medial line: graph representation and shape description, *Proc. 8th Int. Conf. on Pattern Recognition*, Paris, pp. 430–432 (1986).
5. P. P. Cortopassi and T. C. Rearick, A computationally efficient algorithm for shape decomposition, *Proc. 2nd Int. Conf. on Computer Vision*, Ann Arbor, pp. 597–601 (1988).
6. C. Arcelli, R. Colucci and G. Sanniti di Baja, On the description of digital strips, *Proc. Int. Conf. on Artificial Intelligence Applications and Neural Networks*, Zurich, pp. 193–196 (1990).
7. E. Thiel and A. Montanvert, Shape splitting from medial lines using the 3–4 chamfer distance, *Visual Form Analysis and Recognition*, C. Arcelli, L. P. Cordella and G. Sanniti di Baja, eds, pp. 537–546. Plenum Press, New York (1992).
8. H. Blum, A transformation for extracting new descriptors of shape, *Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, ed., pp. 362–380. M.I.T. Press, Cambridge (1967).
9. C. Arcelli and G. Sanniti di Baja, A width-independent fast thinning algorithm, *IEEE Trans. Pattern Anal. Mach. Intell.* **7**, 463–474 (1985).
10. L. Dorst, Pseudo-Euclidean skeletons, *Proc. 8th Int. Conf. on Pattern Recognition*, Paris, pp. 286–288 (1986).
11. C. Arcelli and G. Sanniti di Baja, A one-pass two-operations process to detect the skeletal pixels on the 4-distance transform, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 411–414 (1989).
12. C. Arcelli and M. Frucci, Reversible skeletonization by (5,7,11)-erosion, *Visual Form Analysis and Recognition*, C. Arcelli, L. P. Cordella and G. Sanniti di Baja, eds, pp.21–28. Plenum Press, New York (1992)
13. G. Sanniti di Baja, Well-shaped, stable and reversible skeletons from the (3,4)-distance transform, *J. Visual Commun. Image Repres.* **5**, 107–115 (1994).
14. G. Borgefors, Distance transformations in arbitrary dimensions, *Comput. Vision Graphics Image Process.* **27**, 321–345 (1984).
15. G. Borgefors, Distance transformation in digital images, *Comput. Vision Graphics Image Process.* **34**, 344–371 (1986).
16. G. Borgefors, Another comment on "A note on Distance transformation in digital images", *CVGIP Image Understanding* **54**, 301–306 (1991).
17. E. Thiel and A. Montanvert, Chamfer masks: discrete distance functions, geometrical properties and optimization, *Proc. 11th Int. Conf. on Pattern Recognition*, The Hague, pp. 244–247 (1992).
18. A. Rosenfeld and J. L Pfaltz, Sequential operations in digital picture processing, *J. ACM* **13**, 471–494 (1966).
19. C. Arcelli and G. Sanniti di Baja, Weighted distance transforms: a characterization, *Image Analysis and Processing* II, V. Cantoni *et al.*, eds, pp. 205–211. Plenum Press, New York (1988).
20. C. Arcelli and G. Sanniti di Baja, Finding local maxima in a pseudo Euclidean distance transform, *Comput. Vision Graphics Image Process.* **43**, 361–367 (1988).
21. A. Rosenfeld, Arcs and curves in digital pictures, *J. ACM* **20**, 81–87 (1973).
22. T. Pavlidis, *Structural Pattern Recognition*, Chapter 7. Springer, New York (1977).

About the Author—GABRIELLA SANNITI DI BAJA received the doctoral degree in Physics from the University of Naples, Naples, Italy, in 1973. Since then she has been working at the Istituto di Cibernetica of the National Research Council of Italy, in the position of Director of Research. Her main research activities concern two-dimensional shape representation, decomposition and description. Since 1991, G. Sanniti di Baja has been the chairman of the Education Committee of the International Association for Pattern Recognition.

About the Author—EDOUARD THIEL was born in Saint-Avold, France, in 1968. He received the B.S. in Pure Mathematics from the University of Strasbourg in 1990 and the M.S. in Applied Mathematics in Computational Geometry in 1991. He is currently preparing his Ph.D. thesis with the group TIMC-IMAG of Grenoble, France. His main interests concern theoretical and applicative aspects of discrete distances, shape representation, decomposition and description, and the development of graphic interfaces.