From Primitive Regions to Shape Partition

S.V. Ablameyko°, M. Frucci* and A. Marcelli§

° Institute of Engineering Cybernetics
Academy of Sciences
Surganova st., 6
200072 Minsk, BELARUS

*Istituto di Cibernetica del CNR
Via Toiano, 6
80072 Arco Felice, ITALY

§Dipartimento di Informatica e Sistemistica
Universita' di Napoli "Federico II"
Via Claudio, 21
80125 Napoli, ITALY

Abstract

The work presented in this paper has been developed within the framework of a novel hierarchical method to decompose a figure into a connected set of possibly overlapping components. The method is based on the exploitation of the shape information carried by the skeleton to initially decompose the figure into a number of pattern subsets representing the primitive regions composing the figure. Once the initial decomposition has been obtained, it is exploited to obtain the highest level of our decomposition. This paper addresses specifically the problem of merging the primitive pattern subsets found within the figure to get the minimal number of components. It is assumed that the skeletal pixels are associated with a label specifying the local thickness of the figure, and it is shown that the trend of the label along the skeleton branches allows both to identify the primitive regions within the figure and to merge them into the figure components. Experimental results show that the decomposition provided by the algorithm is very natural in both the number and the type of the components.

1. Introduction

The pivotal role of the shape in both natural and artificial vision systems has been largely recognised, and shape representation and description have been largely addressed in the literature [1, 2]. Typically, a suitable representation of the object is decomposed in such a way that each part or component of the initial representation can be assumed as representing one of the regions constituting the object. Then, the description of the object is obtained in terms of the description of the regions and of their spatial relationships. The decomposition of shape into parts requires the definition of both the parts in which the original shape has to be decomposed and an effective procedure to perform the decomposition. These two aspects of the problem leads to what has been called the shape dilemma: general notions of shape - needed to deal uniformly with the large variability of the visual shapes - tends to be qualitative, while algorithms - needed to compute the properties of interest of specific shapes - tends to focus on details [3].

In a previous paper, we have presented a decomposition algorithm which assumes as initial shape representation the skeleton provided by the algorithm proposed in [4], and provides as result subsets of pixels, each one corresponding to one of the elementary regions composing the figure [5]. The figure decomposition is provided in terms of three primitive regions, called blob, ribbon and bridge, whose definitions and implementation have been described elsewhere [6]. The algorithm is articulated into three main steps. The first step is devoted to select the skeletal pixels which will mainly contribute to the final decomposition. The purpose of the second step is that of ascertaining whether the skeletal pixels selected in the previous step represent the desired decomposition or more pixels are needed, in order to preserve the adjacency relations among the regions. Eventually, the third step of the algorithm is entered for deciding, among the skeletal pixels selected in the second step, which ones represent an actual bridge and those that are rather an extension of one of the regions already detected.

In this paper, after briefly illustrating the definitions of primitive regions we have adopted and the basic features of the decomposition algorithm, we present the criteria we have developed for merging sequences of skeletal pixels corresponding to figure regions in order to get the minimal number of primitive regions needed to provide the desired decomposition.

2. Primitive Regions

We have assumed that the primitive regions composing a shape can be grouped in two categories, depending on their perceptual relevance within the shape under observation.
Regions belonging to the first category convey most of the information associated to the shape, and, in this respect, are essential to characterize the figure. They should be preserved under "small" perturbations in the viewing condition, such as those due to border noise, figure rotation, partial occlusion. They may have a roughly constant thickness or a nearly circular shape, and we have called them either blobs or ribbons. A blob is a region which is adjacent to regions having a smaller thickness, while a ribbon has at least one thicker adjacent region.
On the contrary, regions belonging to the second group represent parts of the shape which are not essential to its description, such as those parts of the figure which may appear or disappear because of any of the factors mentioned before. These regions generally exhibit a thickness that changes monotonically along its main axis. We have called them bridges.

In order to formalize these definitions in terms of the type of information associated to the skeletal pixels, we have introduced the notion of propagating pixel. A propagating pixel is a skeletal pixel belonging to a sequence of pixels whose labels change monotonically along the branch of the skeleton. A directional information associated to the pixels of the sequence specifies the direction along which the labels increase [4]. According to this definition, a blob is defined as a sequence of non-propagating pixels with adjacent propagating skeletal pixels that propagate in opposite directions, towards the interior of the blob. Similarly, a ribbon is a sequence of non-propagating pixels with at least one adjacent propagating skeletal pixels not propagating towards it. Eventually, a bridge is a sequence of skeletal pixels which does not lie within the figure obtained by the union of blobs and ribbons [6]. Each primitive region is obtained as the union of the discs associated to the pixels of a suitable subset of the distance labelled skeleton.

According to the definitions given above, the distinction between ribbon and blob is mainly based on directional information, which depends on the global shape of the pattern at hand. Therefore, the same subpattern may be interpreted as a ribbon or a blob depending on the other subpatterns of the figure, thus reflecting in a quantitative way the changes in shape perception which occur when the same shape is stretched or expanded.

3. Decomposition algorithm

The decomposition algorithm, as mentioned in the introduction, is articulated into three steps. In the first one, the skeleton branches are parsed, the trends of the labels and directional information are analyzed, and the sequences of non-propagating pixels are labelled as blobs or ribbons, depending on the case. These sequences constitute the core of the final decomposition. During the parsing, sequences belonging to either a blob or a ribbon and separated by short sequences of propagating pixels are merged to form longer ones. Sequences which are not labelled as blob or ribbon are left to the next steps for further processing.

The second step performs a reverse distance transformation of the core pixels. If all the skeletal branches lie inside the reconstructed pattern, then the core represents the final decomposition, and the algorithm terminates. If this is not the case, the pixels belonging to these sequences not included in the reverse distance transform are called candidate bridge pixels, and are processed in the last step. To this purpose, the trends of the labels of the candidate bridge and of the adjacent region of the core are analyzed. If they exhibit similar values, the candidate bridge pixels are considered as an extension of the region of the core. Otherwise, only the labels associated to the extremes of the sequence corresponding to the candidate bridge are analyzed. Depending on their values, the candidate bridge pixels are considered as an actual bridge or a ribbon. Note that, to deal with the last case, it is necessary to relax to definition of ribbon given above, by allowing sequences of propagating pixels.

4. Region growing by sequence merging

The definitions of primitive regions and the decomposition algorithm described in the previous sections lead to the following conclusions:

- the core of the decomposition, i.e. the union of blobs and ribbons, should represent the set of primitive regions which are essential to characterize the figure. From this point of view, it should be a kind of "invariant", characterizing the figure at hand, and therefore should be very robust with respect to the different sources of noise which may affect the digital image;
- the bridges, on the contrary, represent "additional" regions of the figure, not essential for its characterization. These regions, for instance, could appear or disappear in rotated instances of the same figure, or in samples affected by different amount of noise;

- the given definitions of primitive regions would lead to a decomposition of the original figure into a larger number of parts, with respect to its "natural" decomposition. In fact, sequences of non propagating pixels interleaved by propagating pixels would be split into a number of pieces, each one corresponding to a sequence of connected non propagating pixels terminated by one (or more) propagating pixel. This would result, in turn, into a number of either blobs or ribbons, largely overlapping and separated by bridges. As a consequence, different cores could be obtained for a given figure, thus contradicting the assumed "invariant" nature of the core itself.

To avoid this inconvenience, it is therefore necessary to discriminate among the candidate bridge pixels which ones must be considered as "noise" of the core, and therefore added to it, and which ones, on the contrary, actually represent an "additional" part of the figure main body, in the sense specified above. To this purpose, we have defined the following set of rules to concatenate sequences of non propagating pixels separated by short sequences of propagating pixels:

1) The region corresponding to a sequence of propagating pixel, called b-sequence, is merged with an adjacent core region if the skeletal sequence obtained after the merge does not exhibit large variation of the labels in correspondence of the merging region.

2) Each b-sequence not satisfying the condition 1) can be regarded as either a bridge or a ribbon. In particular, the b-sequence, at least three pixels long, is a ribbon if the labels along the sequence do not change significantly. To evaluate whether the labels along a sequence change significantly or not, the length of the minimum path between the extreme of the sequence is computed taking into account only the values associated to these extreme. If this length is greater than the smallest weight used to compute the skeleton, the b-sequence is assumed to represent an actual bridge, while it is regarded as a ribbon otherwise.

Illustrations of the results obtained while adopting the rules 1) and 2) are given in fig. 1.

5. Conclusion

In the framework of a novel hierarchical method to decompose a figure into a connected set of possibly overlapping components, we have presented here a set of rules for merging the set of primitive regions detected on a suitable figure skeleton in order to get the simplest decomposition of the figure at hand. The rules, given in terms of both the distance label and directional information associated to the skeletal pixels, greatly contribute to the stability of the final decomposition with respect to not significant changes in the figure shape.

REFERENCES

[1] C. Arcelli, L.P. Cordella and G. Sanniti di Baja eds., *Visual Form: Analysis and Recognition.* New York, NY: Plenum, 1992.

[2] C. Arcelli, L.P. Cordella and G. Sanniti di Baja eds., *Aspects of Visual Form Processing.* Singapore: World Scientific, 1994.

[3] B.B.Kimia, A.Tannenbaum, and S.W.Zucker, "On the shape triangle", in: *Aspects of Visual Form Processing*., C. Arcelli, L.P. Cordella and G. Sanniti di Baja eds., Singapore: World Scientific, pp. 221-230, 1994.

[4] M. Frucci and A. Marcelli, "Efficient skeletonization of binary figures through (d1,d2)-erosion and directional information" in:*Aspects of Visual Form Processing*., C. Arcelli, L.P. Cordella and G. Sanniti di Baja (eds.), Singapore: World Scientific, 1994, pp. 221-230.

[5] S. Ablameyko, M. Frucci, and A. Marcelli, "Shape Decomposition by (d1,d2)-Weighted Skeleton and Directional Information", *Proc. ICPR'96*, Vienna (AUSTRIA), August 22-25, 1996,vol. II, pp. 275-279.

[6] S. Ablameyko, M. Frucci, and A. Marcelli, "Extracting Shape Primitives from Skeleton", *Proc. IAPR Workshop on Machine Vision and Application*, Tokyo (JAPAN), November 13-16, 1996, pp. 502-505.

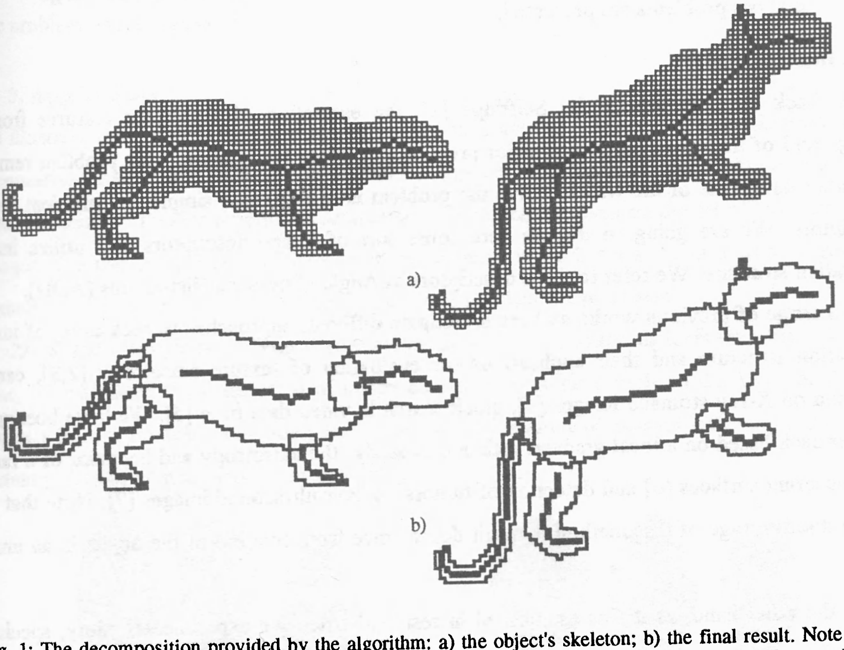

Fig. 1: The decomposition provided by the algorithm: a) the object's skeleton; b) the final result. Note that the disappearance of the panther's ear in the images on the left column of the figure, due to border noise, results in the disappearance of a bridge into the final decomposition, but the core is left untouched.