

# Extracting Shape Primitives from Skeleton

Sergey V. Ablameyko<sup>1</sup>    Maria Frucchi<sup>2</sup>    Angelo Marcelli<sup>3</sup>

<sup>1</sup> Institute of Engineering Cybernetics  
Belarusian Academy of Sciences

<sup>2</sup> Istituto di Cibernetica del CNR

<sup>3</sup> Dipartimento di Informatica e Sistemistica  
Universita' di Napoli "Federico II"

## Abstract

The paper discusses how the set of shape primitives to be used for decomposing a shape into meaningful parts may affect the performance of the decomposition method, and proposes a new set of shape primitives to be used when the input shape is represented by its skeleton. The suggested shape primitives combine the qualitative, general notion of shape needed to cope with the large variability of objects one would like to describe in terms of the same shape, with the unique, computable definition required by shape recognition algorithms. Their definition is given in a context-dependent way, without using any threshold to distinguish among them. The results provided by a decomposition algorithm based on criteria combining both local and global information on the shape of the figure at hand are very natural with respect to both the number and the description of the primitives forming the input shape.

## 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 considered as crucial steps for any successful machine vision task in many different fields, such as document analysis and processing, security systems, robot vision and factory automation [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. In turn, the description of each elementary region is obtained by exploiting the information associated to the corresponding component of the initial object representation.

Independently of the specific technique adopted to obtain it [3-5], the ideal shape description should be able to capture the many dimensions of shape, as well as to provide a way of decomposing the object into its "natural" parts. Furthermore, it should be stable under

rotation and, to some extent, with respect to noise, so that the same description may be obtained from similar objects.

In this paper we show that the set of shape primitives we have adopted provide the flexibility required to deal with the variability exhibited with a large class of similar objects, and, at the same time, allow for a straightforward and cost effective implementation of the decomposition algorithm [6]. We also show that, assuming as initial representation of the object its skeleton computed by using a  $(d_1, d_2)$ -weighted distance transformation [7], so that at every pixel of the skeleton is associated a label specifying its distance from the initial background, the definition of the shape primitives can be given in terms of the information associated to the skeletal pixels. Experimental results are also reported.

## 2 The Shape Primitives

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 lead 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 - tend to focus on details [8].

To solve this dilemma, we have chosen as shape primitives blobs, ribbons and bridges. A blob is an approximately circular-shaped region, while a ribbon is a region of roughly constant thickness, and whose length is appreciably greater than width. Eventually, a bridge is a region whose shape is neither a ribbon nor a blob, but rather an elongated region whose thickness may change monotonically along its main axis. It is called bridge because it mainly represents a part of the shape connecting two blobs, two ribbons, or a blob and a ribbon. It is worth noticing that, in the above definition of blob, circular has to be interpreted with reference to the adopted metric in the digital plane [7].

These definitions have been implemented through the notions of significant and pivot skeletal pixels.

A skeletal pixel  $p$  will be declared significant if one of the following condition holds:

- a) it is an end point;
- b) it has two 8-neighbours and
  - b1) at least one of them has the same label as  $p$ ;
  - b2) all their labels are either greater or smaller than the label of  $p$ ;
  - b3) they have already declared significant;
- c) it is a branch point and
  - c1) there are no significant pixels in its neighbourhood;
  - c2) there are two or more significant pixels in its neighbourhood;

Note that the conditions listed under letter b) allow for small variations of the labels of the pixels within a sequence, thus implementing the notion of roughly constant thickness mentioned before, without using any threshold. Once the significant pixels of the skeleton have been detected, to the other ones are associated pointers pointing to the neighbour pixel whose label is the largest of the neighbourhood.

A significant pixel  $p$  will be declared pivot if one of the following condition holds:

- a)  $p$  has no neighbours;
- b)  $p$  belongs to a sequence of at least three significant pixels;
- c)  $p$  belongs to a sequence of significant pixels and at least one of the adjacent pixels has a pointer directed toward the sequence.
- d)  $p$  is a branch point declared significant under condition c1

By using the notion of pivot pixel, the shape primitives can be defined as follows. A sequence of skeletal pixels represents a blob if one of the following conditions holds:

- 1) it is made by just one pivot pixel;
- 2) it is a sequence of pivot pixels whose adjacent skeletal pixels have pointers pointing to it.
- 3) it is a sequence of pivot pixels delimited by two end points.

A sequence of at least three skeletal pixels represents a ribbon if one of the following conditions holds:

- 1) it is a sequence of pivot pixels whose adjacent skeletal pixels have pointers with the same orientation;
- 2) it is a sequence of significant pixels, whose adjacent skeletal pixels have pointers with the opposite orientation not directed toward the sequence.

These conditions show that, by using only local information, namely the labels, the skeletal pixels are divided in two classes: pixels which represent regions of the figure of quasi-constant thickness, i.e., pixels whose labels have similar values, and the others. The first one are eligible to represent ribbon and/or blobs, and therefore to be included in the set of pixels constituting the final decomposition. To decide whether these pixels will be actually included in the final decomposition and, in the affirmative, whether they represent a blobs or ribbons, local information does not suffice, and therefore global information on the shape of the object is gathered and used. This global

information is obtained by looking at the trend of the labels associated to the pixels of the second class belonging to the same skeletal branches to which the pixels of the first class belong to. Fig. 1a shows one of these cases, where local information, namely the labels of the sequence of pixels labelled 12, would lead to consider the sequence as a ribbon, and therefore to include the pixels of the sequence in the final decomposition. On the contrary, global information, shows that the region associated to the sequence does not actually correspond to a blob or a ribbon, and therefore should not included in the final decomposition. The same problem arises to decide whether the pixels of the second class represent a bridge, and therefore have to be included into the final decomposition, or may be deleted. To tackle the problem we refer to the notion of power of expansion of the pixels already included in the final decomposition. Given these pixels, the associated figure can be reconstructed by means of a reverse distance transformation applied to them. It may happen that some skeletal pixels not included in the final decomposition fall outside the reconstructed figure, and therefore they candidate to represent bridges or ribbons, according to our definition. If the skeletal pixels not included into the reconstructed figure belong to the first class, they are added to the final decomposition and considered as representing a ribbon. In case the pixels not included into the reconstructed figure belong to the second class, they are considered as candidate bridge pixels, and to decide whether they actually represent a bridge, the skeletal pixels belonging to the same branch as the candidate bridge and lying inside the reconstructed figure are expanded, i.e. a reverse distance transformation is applied to them. If the resulting reconstructed figure does not contain all the pixels of the branch, the candidate bridge pixels are considered as representing a bridge and included into the final decomposition. Otherwise, the candidate bridge pixels, as well as the skeletal pixels connecting them to the smallest region between the two associated to the branch, are added to the final decomposition.

### 3 Discussion and Conclusion

In this paper we have proposed a set of shape primitives and a shape decomposition method which may represent a solution to the generic shape dilemma mentioned earlier, in that our shape primitives seem to be general enough to cope with the large variability of visual shapes, while providing, at the same time, the unique description required by shape analysis algorithms.

Fig. 1 shows one of the main features of the proposed set of shape primitives: the pixels belonging to the same sequence, namely the sequence of pixels labelled 12, are considered as forming a blob, a ribbon or a bridge according to the context, i.e., the shape of the whole object they are part of. This is achieved by looking at both the shape of the regions and how they

are combined into the object. Fig. 2 shows that the adopted primitives lead to a natural - with respect to both the number and the type of the primitives forming the input shape - and rotation invariant decomposition.

## 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] T. Pavlidis, "A Review of Algorithms for Shape Analysis", *CGIP*, vol. 7, pp. 243-258, 1978.
- [4] S. Marshall, "Review of shape coding techniques", *Image and Vision Computing*, vol. 7, no. 4, pp. 281-294, 1989.
- [5] C. Teh and R.T. Chin, "On the detection of dominant points on digital curves", *IEEE Trans. on PAMI*, vol. 11, pp. 859-872, 1989.
- [6] S. Ablameyko, M; Frucci, and A. Marcelli, "Shape Decomposition by (d1,d2)-Weighted Skeleton and Directional Information", *Proc. ICPR'96*, vol.II, pp. 275-279.
- [7] G.Borgerfors, "Distance transformations in digital images", *CVGIP*, vol. 34, pp. 344-371, 1986.
- [8] 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.

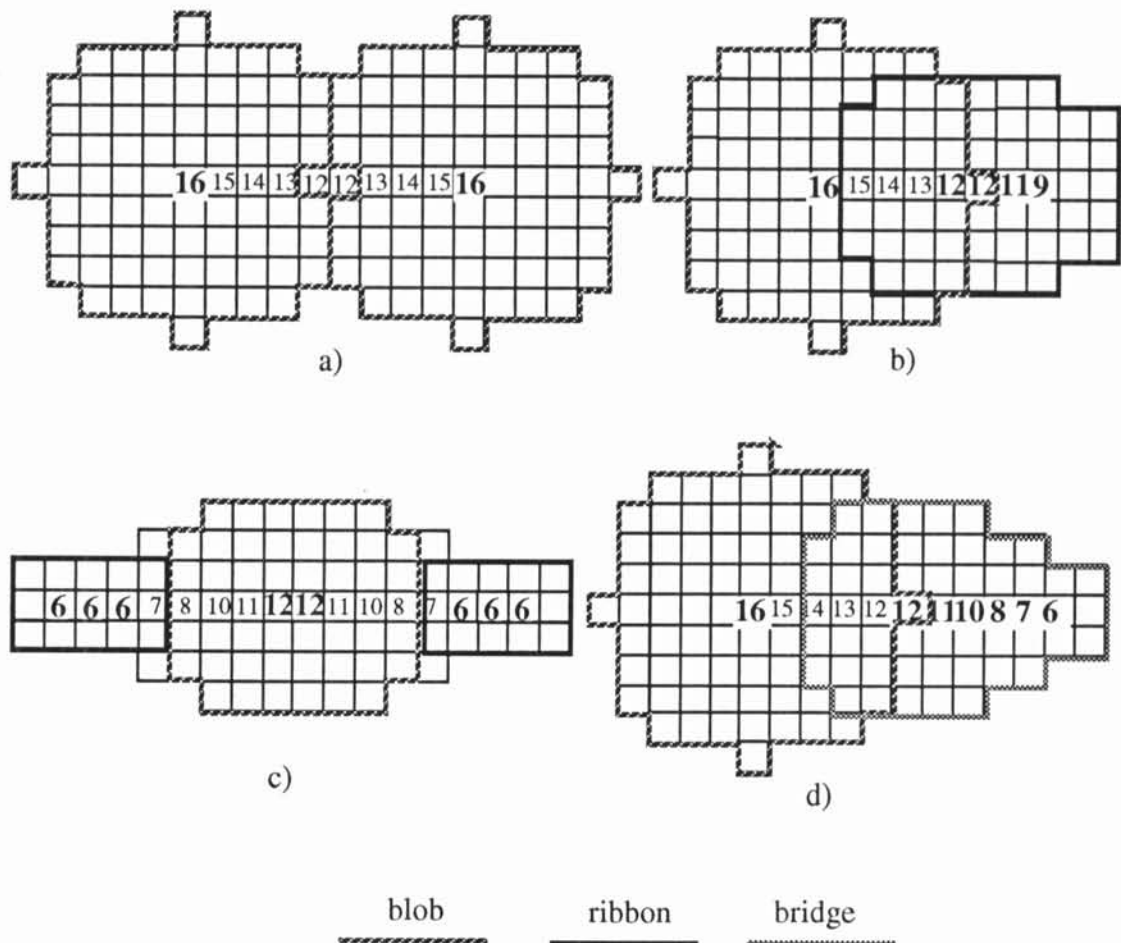
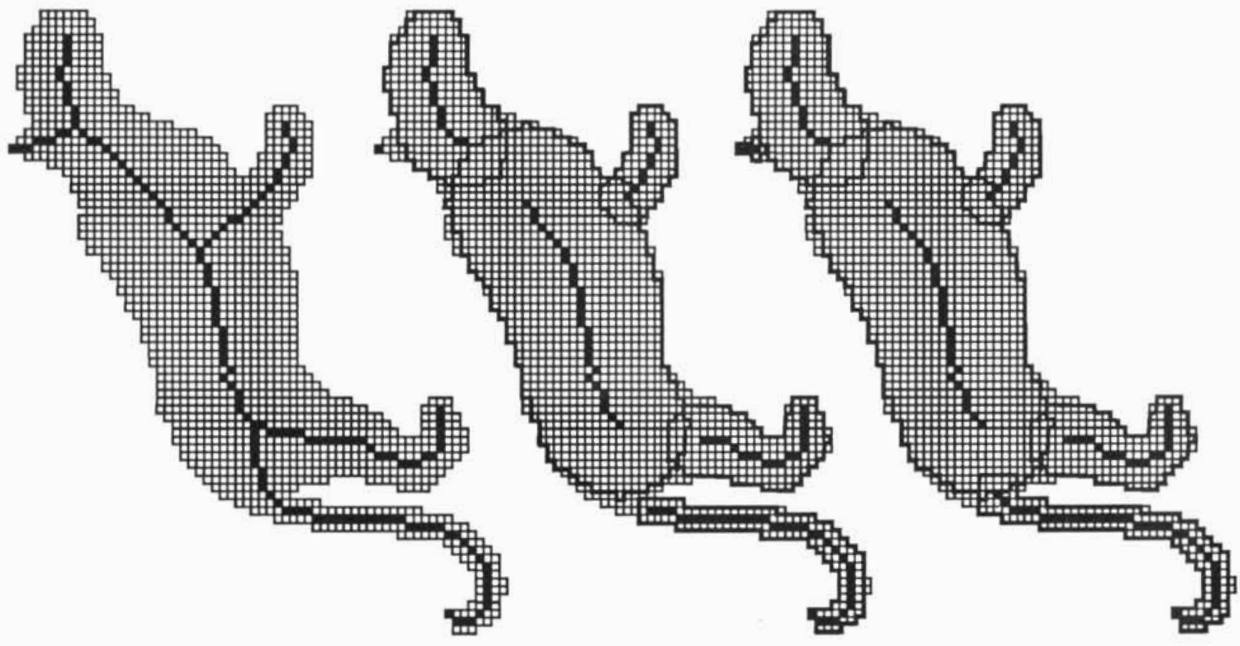


Fig. 1. Illustration of the shape primitives. The skeleton has been obtained by adopting a (3,4)-weighted distance. The set of sequences of adjacent skeletal pixels in boldface represent the skeleton decomposition. The regions obtained by applying the reverse distance transformation to the sequences are the components of the pattern.. The same sequence of pixels of the first class with label 12 is not included in the final decomposition (a), as belonging to a ribbon (b), is labeled as blob (c), part of it is labelled as bridge (d), depending on the global shape of the object.



a)

b)

c)

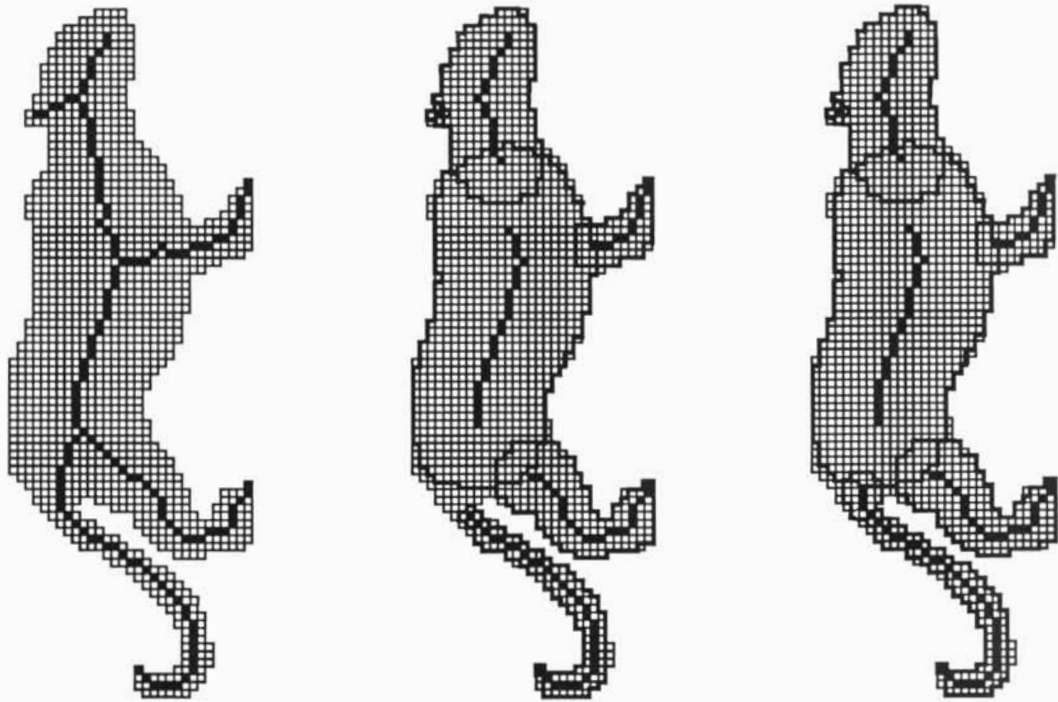


Fig. 2. The results of the decomposition algorithm applied to rotated figures: a) skeletons; b) intermediate results; c) final decompositions.