Skeletonization of Gray-Tone Images Based on Region Analysis

Luca Serino

Istituto di Cibernetica "E. Caianiello", CNR, 80078 Pozzuoli (Napoli), Italy
ls@imagm.cib.na.cnr.it

Abstract. A problem often present in skeletonization of gray-tone digital images is that the obtained skeleton includes an excessive number of branches. In this respect, a regularization process should be performed in order to partially, or totally, remove branches which are not meaningful in the problem domain. In this paper, we propose a skeletonization algorithm which is active only on a suitable subset of the image, mainly constituted by regions understood as relevant from a perceptual point of view. The notion of dominance of a region, which is defined in terms of geometrical features, gray-value and adjacency relations, plays a central role in the selection of the regions of the subset. The obtained skeleton turns out to be more representative and its simpler structure will allow one to perform the regularization phase with a reduced computational effort.

1 Introduction

Generally, the skeleton of a gray-tone digital image could be defined as a set having a "graph-like" structure which, ideally, should represent a sketch of the subset of the image which is understood as the foreground in a specific problem domain. However, since it is often difficult to distinguish clearly which regions belong to the foreground, the obtained skeleton may include branches found in correspondence with non significant regions, and its structure may result too busy for representation purposes. In this respect, the representation power of the skeleton can be improved noticeably by taking into account a regularization process, which modifies the skeleton structure by using context information and domain knowledge [1], [2], [3].

In this paper, we propose a skeletonization algorithm which is active only on a suitable subset of the image, mainly constituted by regions understood as relevant from a perceptual point of view. Purpose of this selection of the image subset is to highlight the regions which are more likely to belong to the foreground. The advantages inherent in this approach with respect to previous work, e.g., [4], are: i) a reduced computational effort during skeletonization, since the skeleton is searched only in a subset of the image. ii) a skeleton having a simpler structure and therefore more representative. iii) a reduced computational effort during the regularization phase, since the skeleton is likely to be constituted by a smaller number of branches.

To find the regions belonging to the image subset of interest, we introduce the notion of dominant region. This notion involves geometrical features, gray-value and adjacency relations of a region, and turns out to be useful in classifying a region as perceptually more relevant than (some of) its adjacent regions. The image subset will consist of the dominant regions and of a number of other regions, called induced regions. The induced regions, although not perceptually meaningful, are important because are placed along an ideal path connecting a dominant region with another

A. Sanfeliu et al. (Eds.): CIARP 2004, LNCS 3287, pp. 495–502, 2004.

dominant region with higher gray-value, and their presence allow connectedness preservation during the pixel removal phase.

The proposed skeletonization algorithm follows the scheme described in a previous paper [4], which allows a computationally convenient analysis of the regions to be processed. Besides the region selection phase, it includes a labeling phase guided by distance transformation [5], [6], [7], and a pixel removal phase accomplished by topology preserving reduction operations [8], [9].

2 Preliminaries

Let *G* be a gray-tone digital image. Pixels in *G* are assigned one out of a finite number of increasing integer values g_k k=0,1,...,N, which indicates for any pixel *p* the gray-value or status *g(p)* of the pixel itself. We assume that *G* is bordered by a frame of pixels with gray-value higher than g_N. In the following, pixels will be described as darker as their gray-value is higher.

The neighbors of *p* are its 8-adjacent pixels. They constitute the neighborhood *N(p)* of *p* and are denoted by $n_1,n_2,...,n_8$, where the sub indexes increase clockwise from the pixel n_1 placed to the left of *p*.

A gray-tone digital image is a mosaic constituted by regions, which we regard as maximal 4-connected sets of pixels with the same gray-value.

Two regions in the mosaic are adjacent if they are 4-adjacent.

The area of a region is the number of its pixels.

The perimeter of a region is the number of its pixels 4-adjacent to the adjacent regions.

The length of the common border between a region and an adjacent one is the number of its pixels 4-adjacent to that region.

A “higher neighboring component” (shortly, HNC) of a region *R* is a maximal connected set of adjacent regions, each having gray-value higher than the one of *R*.

A “lower neighboring component” (shortly, LNC) of a region *R* is a maximal connected set of adjacent regions, each having gray-value lower than the one of *R*.

3 Regions

In a gray-tone image where there is no *a priori* knowledge of its contents, we consider the darker areas as foreground and those clearer as background. To regard a region as darker does not depend on its real gray-value, but only on the existence of neighboring regions with lower gray-value. In a broad sense, darker regions are perceptually more relevant and, under certain conditions, we say that they dominate the neighboring regions.

In order to detect the dominant regions, it is preliminarily convenient to distinguish four types of regions: top, bottom, saddle and slope.

- A top is a region with gray-value higher than the gray-value of all its adjacent regions.
- A bottom is a region with gray-value lower than the gray-value of all its adjacent regions.
- A saddle is a region for which there exist at least either two HNCs or two LNCs.
- A slope is a region for which there exists exactly one HNC and one LNC.

Since the perceptual relevance of a region depends on the local context in which it is placed, the tops assume a determinant role since they are characterized by a locally maximum gray-value. Thus, all tops should be represented by branches of the skeleton. On the contrary, the bottoms have a locally minimum gray-value and are very likely to be part of the background. Such regions should not be represented by branches of the skeleton.

More questionable is the decision regarding saddles and slopes.

A saddle is a region that separates two HNCs (or two LNCs) and may correspond either to an abrupt change of gray-value with respect to the gray-values of these components or to a smooth transition between those gray-values. Whichever the case, we regard a saddle as part of the background if its gray-value is closer to the gray-values of the LNCs than to the ones of the HNCs, and as part of the foreground if this is not true, i.e., the saddle appears sharply defined. Thus, a saddle cannot be considered automatically as a dominant region, and certain measurements have to be performed to this purpose. See Fig. 1 (a,b).

Slopes correspond to perceptually relevant regions if they are elongated and mostly surrounded by regions with a suitably lower gray-value, i.e., are sharply defined. In this case, they are dominant regions and skeleton branches should be found in correspondence with them. In other cases, they could be understood as belonging either to the foreground or to the background, depending on the context. Particularly, we regard them as part of the foreground if they are useful (to contribute) to establish a connection between dominant regions with different gray-values.

To define the set of the dominant regions, we propose the four criteria below, which take into account gray-value, elongation, and local context.

c1: the region is not a bottom;
c2: the region is a top;
c3: the region is a sharp saddle;
c4: the region is an elongated and sharp slope;

Then, we say that a region is dominant if the following condition is verified:

c1 AND (c2 OR c3 OR c4)

Once the dominant regions have been found, it is important to establish which other regions can be understood as belonging to the foreground, and whose detection can be induced by the presence of the dominant regions. For instance, if we identify an elongated slope as a dominant region, then it is important that it be connected to the part of the foreground (i.e., already detected dominant regions) of which it is perceived as a protrusion. To this purpose, it is necessary to consider as belonging to the foreground also the slopes placed between the protrusion and the foreground. Once detected, an induced region will induce in turn other regions with higher and higher gray-values until a dominant region is found. See Fig. 1c.

4 Skeletonization

The skeletonization process can schematically be divided in 5 phases. The first phase regards the preprocessing. The aim is to simplify the input image by using an iterative merging process, which creates macro regions including input regions whose gray-values can vary only within a given range [10]. We don't discuss this phase since it is described in detail in a previous paper [4]. The second phase regards the extraction of

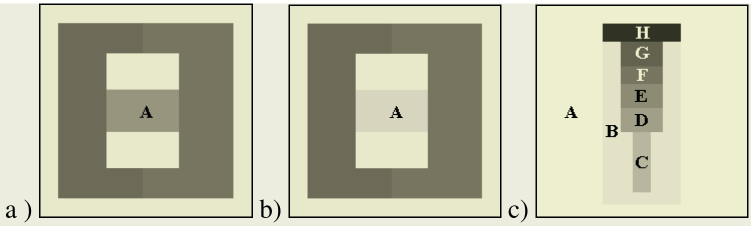


Fig. 1. a-b) Region A is a saddle. Depending on the context, A can be dominant (a) or not (b). c) Region A is a bottom. Region H is a top. Regions B-G are slopes. Slopes D, E, F, and G are induced from the elongated and sharp slope C.

data characterizing the regions. The third phase regards the detection of both the dominant regions and the regions induced by these ones. The fourth phase regards the extraction of the skeleton from the regions previously selected. We will only briefly outline this phase, since it is not significantly different from a similar one described in detail in [4]. The last phase regards the post processing, tailored to originate a one-pixel-thick skeleton.

4.1 Data Characterizing Regions

During this phase we collect useful data regarding the regions. In order to keep these data, we use a list *R* of records; every record of *R* has 10 fields and contains the data concerning a region. The first field is a natural number that is ascribed to the region and has the role of ID of the region; the fields from 2 to 5 contain respectively the gray-value, the area, the length of the common border with the regions having lower gray-value, the length of the common border with the regions having higher gray-value; the fields 6 and 7 contain numbers 0, 1 or 2 depending on whether the number of HNCs (LNCs) is = 0 or = 1 or > 1; the fields 8 and 9 points to two other lists of records: the first one points to a list containing the ID of the adjacent regions of the examined region; the second points to a list *C* of records keeping the coordinates of the pixels of the region. Finally the last field is a Boolean field that will assume true or false value depending on whether the region will be considered as foreground or background.

Since it is important to access the data regarding a region also by starting from the ID value, we have chosen as ID the progressive numbers starting from 1. In this way, we can use an array of pointers in which the generic element “i” points to the record containing the data of the i-th region.

The extraction from the image *G* of the data of interest regarding every region is carried out in two steps, and in different ways.

The first step is concerned with a raster scan of the image *G* during which, every time that a new region is detected, the progressive number is ascribed as ID to the region, and a new record is created where the ID and the gray-value of the region are inserted. All the pixels of the region are then detected and labeled with their ID in a new array *F*. While the pixels of a region are detected, the border pixels regarding the region itself are inserted into the list *C* and the value of the area is also computed. Once the entire region has been examined, also this value will be inserted in the relative field of the record.

The second step is concerned with the analysis of the list R, in order to assign to every region the ID of the adjacent regions and the lengths of the common border with its adjacent regions. To this purpose, we consider, for every region, the list C of the border pixels. Every member of C is a pair of coordinates characterizing a pixel in F. The set of the regions 4-adjacent to the border pixels identifies the regions adjacent to the examined region. Every time that a new adjacent region is found, its ID is inserted in the list of the region under examination. By analyzing the 4-neighbors of the pixel in F it is also possible to know the gray-value of the adjacent regions. Therefore, after analyzing the list C, the length of the common border with the regions with higher and lower gray-values will be known and these lengths can be inserted in the relative fields of the record.

The R list is analyzed again in order to detect, for every region, the HNCs and the LNCs; in particular, it is necessary to know the number (0, 1, >1) of HNCs and LNCs. First, we consider the case of the HNCs. The description is referred to a generic region found when scanning the list R, since the procedure is the same for all the regions. We consider the list of the adjacent regions and construct a new list L of regions including only the regions with gray-value higher than the one of the examined region. If L is empty, the number of HNCs is equal to zero, else we analyze L in order to check, by considering the adjacency of the regions stored in L, whether all these regions constitute only one connected set in the image F. To this purpose, we consider a new list L', initially empty. We remove the first element from L and place it in L'. For every region E of L, we consider its list of adjacent regions in order to detect whether an element of L' is in it. If this happens, we remove the region E from L and insert it in L'. When the analysis of all the elements of L terminates, one of the following three cases can occur. i) The list L is empty; in this case the procedure terminates and the number of components found is equal to 1. ii) The list L is not empty, but at least one removal has been done; in this case L must be scanned again. iii) The list L is not empty and no removal has been done; in this case the procedure terminates and the number of components found is greater of 1.

The procedure to detect the LNCs is analogous; the only difference occurs in the construction of L, where the elements with lower gray-value will be inserted.

4.2 Selection of Dominant and Induced Regions

The conditions c1, c2, c3, c4 mentioned above (see section 3) are taken into account to decide whether a region is dominant.

The first condition says that the region should not be a bottom. This information is available since if a region is a bottom, the number of its LNCs is equal to 0.

The second condition says that the region should be a top. Also this information is available since a top has the number of HNCs equal to 0.

The third condition says that the region should be a sharp saddle. If the region is a saddle the number of HNCs (LNCs) is equal at least to 2. If this is the case, it is necessary to check whether this region is sharp. To this purpose, it is necessary to have a measure for the "nearness" of the saddle with respect to the HNCs and with respect to the LNCs. In our opinion, a crucial role in defining the quality of the context for the HNCs (LNCs) is played both by the gray-value of the regions that constitute the HNCs (LNCs) and by their spatial extensions. Thus, we ascribe to the HNCs (LNCs) a weight equal to the sum of the products between the gray-value and the area of every region of the HNCs (LNCs), divided by the area of the HNCs (LNCs). We indicate such values with *d_hnc* and *d_lnc* respectively. If there results that *d_hnc*<*d_lnc*, that is if the gray-value of the region is nearer to *d_hnc* than to *d_lnc*, we take the region as a dominant region. In order to evaluate *d_hnc* and *d_lnc*, only the data regarding the gray-values and the areas of the adjacent regions are necessary; thus one scan of the list of the adjacencies of the considered region is sufficient.

The fourth condition says that the region should be a sharp and elongated slope, understood as a protrusion. We note that, if the region is a slope, the number of HNCs and of LNCs must be equal to 1, therefore this information is available. With regard to the sharpness, we follow the procedure used for the saddles. What remains to evaluate, is whether the slope is a protrusion. Our criterion is to estimate the ratio between the respective lengths of the common border with LNC and with HNC. If this ratio is greater than a threshold, the slope is a protrusion. In this paper, we have chosen a threshold value equal to 3. This ratio is readily available since the data regarding the lengths of the common borders are already known.

To select the dominant regions is therefore sufficient one scan of the list *R*, then a true or false value will be inserted depending on whether the region is dominant or not.

Once the dominant regions have been detected, the induced regions should be found. In fact, every dominant region induces the adjacent regions with higher gray-value to become themselves dominant regions.

In order to find the induced regions, it is necessary to consider a process starting from the detected dominant regions. Thus, we perform one more scan of the list *R* during which, every time that we detect a dominant region, we begin an iterative process. This process starts from the dominant region and analyzes step by step the regions with higher gray-values adjacent to the created induced regions, and proceeds until an already detected dominant region (possibly, a top) is found. In Fig. 2, a preprocessed input image is shown (a), together with the extracted dominant and induced regions (b).

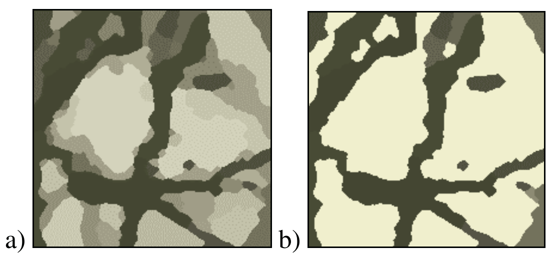


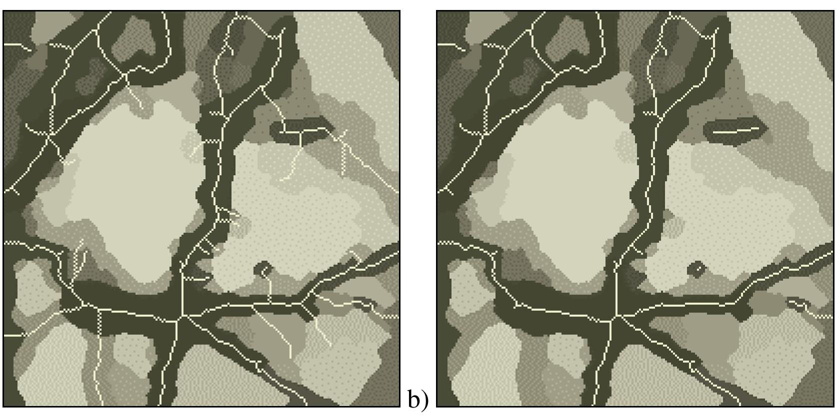
Fig. 2. a) Preprocessed input image. b) Dominant and induced regions.

4.3 Skeleton Extraction and Post Processing

The dominant regions and the induced regions are the only regions involved in the skeletonization process. The skeletonization procedure that we use was introduced in [4]. There, the skeleton was computed by applying topology preserving reduction operations to the pixels of the image, which is analyzed region by region. We outline below the modifications to bring to that algorithm in order to allow its use in the framework of this paper. The algorithm described in [4] is performed in three phases: region labeling, end point detection and pixel removal. Regarding the region labeling phase, this is replaced by the phases described in subsections 4.1 and 4.2 to which one has to add, with the same modalities described in [4], the computation of the distance transform for every region. Regarding the end point detection phase, no change is needed. As for the pixel removal phase, the only modification consists in estimating, every time that a new region is considered, if the region is of interest; if this is not the case, its analysis is skipped and the process continues on the remaining regions. The set obtained at the end of this phase is not ensured to be one-pixel-thick, so that a post processing phase is required. The image is considered as a binary image where the skeleton constitutes the foreground and its complement the background, so it is sufficient to apply well-known topology preserving reduction operations designed for binary images [8] to the set obtained above, during one raster scan of the image.

5 Conclusions

In this paper we have proposed a skeletonization algorithm for gray-tone images, based on region analysis. The aim of this analysis has been to try to foresee which regions could reasonably be understood as belonging to the foreground, so as to exclude the remaining ones from the skeletonization process. We have characterized four types of regions, by taking into account the gray-values of the regions and their spatial relations with the adjacent regions. We have also proposed some criteria to highlight a number of regions, called dominant regions, that are perceptually relevant and should be part of the foreground. The process of region selection foresees the detection of further regions, called induced regions, which allow one to link to each other the dominant regions. The dominant regions and the induced regions constitute the image subset on which the skeletonization algorithm is applied. The resulting skeleton (see Fig.3) is more representative than a skeleton obtained by considering all the regions of the image. Moreover, its simpler structure allows one to implement with a reduced computational effort the regularization phase, which should necessarily be considered to obtain a skeleton meaningful in the problem domain.



a) b)

Fig. 3. Skeleton, superimposed on the preprocessed input, computed by considering all the regions (a) and by considering only the dominant and induced regions (b).

References

1. G. Bertrand, J.-Ch. Everat, M Couprie, Image segmentation through operators based on topology, J. Electronic Imaging **6** (1997) 395-405.
2. C. Arcelli, L. Serino, Regularization of graphlike sets in gray-tone digital images, Int. J. Pattern Recognition and Artificial Intelligence **15** (2001) 643-657.
3. L. Najman, M. Schmitt, Geodesic saliency of watershed contours and hierarchical segmentation, IEEE Trans. on PAMI **18** (1996) 1163-1173.
4. C. Arcelli, L. Serino, Topology preservation and tricky patterns in gray-tone images, in: I. Nyström, G. Sanniti di Baja, S. Svensson (Eds.), Discrete Geometry for Computer Imagery, LNCS 2886, Springer, Berlin, 2003, pp. 298-307.
5. A. Rosenfeld, On connectivity properties of greyscale pictures, Pattern Recognition **16** (1983) 47-50.
6. J. Piper, E. Granum, Computing distance transformations in convex and non-convex domains, Pattern Recognition **20** (1987) 599-615.
7. G. Borgefors, Distance transformations in digital images, Computer Vision, Graphics and Image Processing **34** (1986) 344-371.
8. C. Arcelli, G. Sanniti di Baja, Skeletons of planar patterns, in: T.Y. Kong, A. Rosenfeld, (Eds), Topological Algorithms for Digital Image Processing. North Holland, Amsterdam, 1996, pp. 99-143.
9. V. Ranwez, P. Soille, Order independent homotopic thinning for binary and grey tone anchored skeletons, Pattern Recognition Letters **23** (2002) 687-702.
10. Y. Wang, P. Bhattacharya, On parameter-dependent connected components of gray images, Pattern Recognition **29** (1996) 1359-1368.