3—16

Detection of Curvilinear Structures Using the Euclidean Distance Transform

Jeong-Hun Jang and Ki-Sang Hong

Image Information Processing Lab., Dept. of E.E., POSTECH, Korea

E-mail : {jeonghun,hongks}@postech.ac.kr

Abstract

In this paper, we present a new method for detecting curvilinear structures in a gray-scale image. The concept of skeleton extraction is introduced to detect more general structures such as tapering structures. A skeleton is extracted from the Euclidean distance map that is constructed based on the edge map of an input image. Then, skeletal points are classified into three types (RIDGE, RAVINE and STAIR), and connected points belonging to the same type are grouped to form a skeletal segment. Our detector satisfies many of desirable properties required of a curvilinear structure detector, and moreover it overcomes some limitations of conventional approaches.

1 Introduction

The term *curvilinear structure* denotes a line or a curve with some *width*. Curvilinear structures can be found in most natural images, but their detection is especially useful, for example, when trying to find roads or rivers in aerial images, blood vessels or bones in medical images, and characters in text images.

There are many publications addressing the problem of curvilinear structure detection. Most of recently proposed methods are based on one of following three approaches, and some modifications or new ideas are added to overcome inherent limitations of each approach : 1. Locally parallel edge based approach [1], 2. Ridge based differential geometric approach [2], 3. Active contour model based approach [3].

Since each approach has its own strong and weak points, it is an user's responsibility to choose a method appropriate to the given situation. However, above approaches have a fundamental limitation. Since they are designed to detect only elongated structures with small variation of width along their center lines, they are not adequate for the detection or description of more general structures such as tapering structures. For example, in Approach 2, since the centers of a bar-shaped structure are flat, a Gaussian filter is applied to make them convex. The size of the filter kernel, σ should be large enough to detect wide curvilinear structures, but, then thin lines or curves are blurred out. Therefore, the detectable range of structure's width is confined by the given value of σ. We could find the structures of different width by applying a detector repeatedly with different scales, but it is very costly, and it is not easy to integrate the detection results.

To solve the above problem, we bring the *skeleton extraction* concept, which is famous in the binary image domain, to the gray-scale image domain. That is, what we are trying to do in this paper is to extract a skeleton from a gray-scale image which describes meaningful structures contained in the image. In this approach, as will be shown later, a curvilinear structure detection problem becomes a sub-problem of classifying skeletal segments according to their property. The proposed detector satisfies many of desirable properties required of a curvilinear structure detector, and an extracted skeleton is adequate for describing more complex structures.

A skeleton of some region strongly depends on the boundary of the region according to its definition. Since boundaries of regions are not defined clearly in the case of a gray-scale image, some decision rule should be involved to determine them. We make use of detected edgels (edge pixels) as the alternative to boundary points though they are imperfect. Once boundary points are determined, we can utilize powerful tools which have been developed for skeleton extraction in the binary image domain for many years. There are several approaches for extracting a skeleton. But, all the approaches are not suitable to our situation. For various reasons that will be explained later, we take an approach in which the *Euclidean distance transform* is performed on an edge map, and ridge points (or local maxima), which are regarded as candidate skeletal points, are extracted from the constructed Euclidean distance map. Extracted ridge points usually contain many unnecessary ones which are caused by noisy boundary shapes, incomplete edges, etc. In order to remove them, we combine several, previously proposed and our own methods effectively. Remaining ridge points constitute a skeleton.

Each skeletal point is assigned one of three types (RIDGE, RAVINE, and STAIR) by observing the cross-sectional shape of an image in the neighborhood of the point. Then, connected points of the same type are grouped to form a skeletal segment with the single type label. Unstably labeled skeletal points are reclassified depending on their neighboring segments' types. This classification procedure, together with the Euclidean distance map, enables us to find curvilinear structures with the selected width.

2 Skeleton Extraction

In this section, skeleton extraction in a gray-scale image is considered in detail. The skeleton extraction procedure is composed of many steps, and its overall flow is shown in Figure 1.

Step 1. detecting edgels : The skeleton extraction procedure begins with the detection of edgels in an input gray-scale image. Detected edgels serve as boundary points of meaningful structures. In our work, the *Canny edge detector* is used because of its many desirable properties such as optimality on singnal-to-noise ratio and localization, scale-space representation, and so on [4]. Canny's method consists mainly of four parts : Gaussian smoothing, gradient computation, nonmaxima suppression, and hysteresis thresholding. Three parameters are involved in the algorithm : the size of a Gaussian kernel, σ in Gaussian smoothing, and two thresholds, T_l and T_h in hysteresis thresholding. After edgels are detected, isolated edgels are removed.

Step 2. constructing an Euclidean distance map : We adopted the "medial axis extraction from a distance map" approach among several existing skeletonization approaches, because it was considered to be most adequate to our situation. For example, it does not require closed boundary contours, and it provides the width of a detected structure. We use the *region growing Euclidean distance transform* algorithm proposed by Cuisenaire [5]. It yields as a result an Euclidean distance map where each pixel site has the distance value to the nearest edgel,
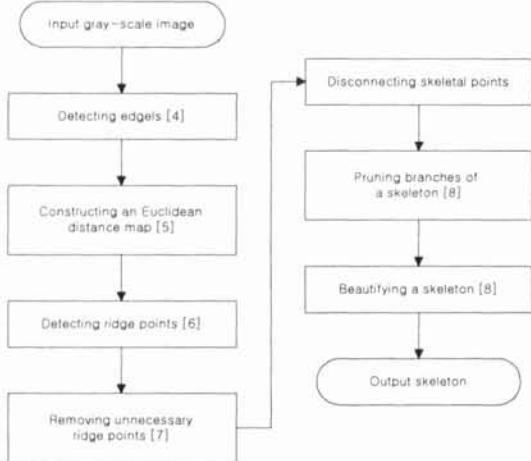
Figure 1: Skeleton extraction procedure.

and the position of the edgel. The algorithm starts computing a distance at each edgel, and the computation region grows outward. Since former results are used to calculate a distance at a new pixel site, the algorithm is fast and efficient.

Step 3. detecting ridge points : Ridge points or local maxima of a distance map constitute candidate skeletal points. We implemented the algorithm proposed by Arcelli and Baja which is specially designed to detect ridges in a distance map [6]. One of advantages of this method is to guarantee connectedness of ridges. In the algorithm, *strong ridge points* are searched for in the raster scan manner, and for each strong ridge point which has been met, *weak ridge points* are tracked starting from it. Strong ridge points can be found using eight predefined operators. The tracking of weak ridge points is based on gradient computation between two neighboring points. Extracted ridges are thinned to be one pixel wide.

Step 4. removing unnecessary ridge points : Step 3 results in a superfluous number of ridge points due to noisy boundaries. Therefore, it needs to remove ridge points that make little contribution to the description of meaningful structures. Malandain and Fernández-Vidal introduced two parameters ϕ and d for the local characterization of skeletal points [7]. The meaning of ϕ and d can be understood easily in Figure 2, where **p** and **n** represent a skeletal point (i.e. ridge point) and its neighbor respectively, and $\mathbf{e}_p$ and $\mathbf{e}_n$ are the corresponding nearest boundary points (i.e. edgels). Parameters ϕ and d of a skeletal point **p** are given by

$$\phi(\mathbf{p}) = \max_{\mathbf{n} \in N(\mathbf{p})} \frac{180}{\pi} \arccos \frac{\mathbf{p}\vec{\mathbf{e}}_p \cdot \mathbf{p}\vec{\mathbf{e}}_n}{\|\mathbf{p}\vec{\mathbf{e}}_p\| \|\mathbf{p}\vec{\mathbf{e}}_n\|}, \quad (1)$$

$$d(\mathbf{p}) = \max_{\mathbf{n} \in N(\mathbf{p})} \|\mathbf{e}_p\vec{\mathbf{e}}_n\| \quad (2)$$

where $N(\mathbf{p})$ denotes an eight-connected neighborhood of **p**. Note that in general, the value of ϕ is small at a point which lies on
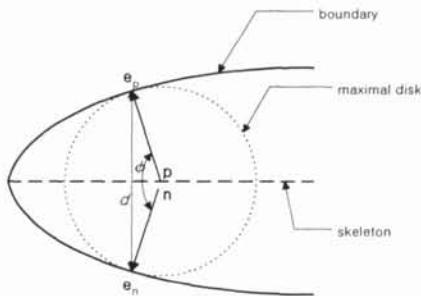

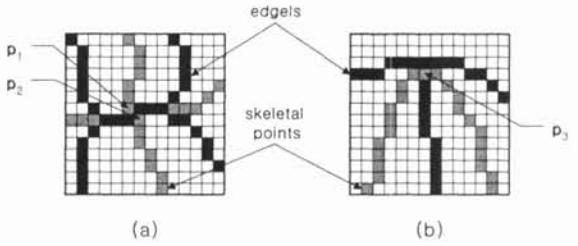Figure 2: Definition of two parameters ϕ and d.

Figure 3: Problems caused by the use of edgels as boundary points.

a skeletal branch generated by small protrusions of boundaries, compared to that of a point on a main skeleton. Simple thresholding of the parameter ϕ is insufficient to remove unnecessary ridge points effectively, since it does not preserve the important topological structure of an original shape. To solve the problem, the *two thresholds* scheme is adopted where two thresholds ϕ_h and ϕ_l ($\phi_h > \phi_l$) are used to produce two skeletons S_h and S_l ($S_h \subset S_l$) respectively. Based on S_h and S_l, a method called *topological reconstruction* is performed to get a robust skeleton [7].

Step 5. disconnecting skeletal points : It is easy to understand that points in a skeleton can be categorized into three kinds of points (i.e. end, link, and junction points) according to their role in a set of connected points. And, it can be shown that they can be identified by observing their eight neighboring points. In our case, however, identifying the points without taking edgels into account causes problems depicted in Figure 3. Two skeletal points $\mathbf{p}_1$ and $\mathbf{p}_2$ in Figure 3(a) are connected to each other, considering an eight-connected neighborhood. But, they should be regared as disconnected, since they are crossing an edge which serves as a boundary. Let S and E denote skeletal and edge point sets respectively, and let $P = S \cup E$. Points to be disconnected can be found in P using two operators in Figure 4, where **p** represents a skeletal point under consideration, and **n** and **e** designate neighboring skeletal and edge points respectively. P is scanned sequentially, and if at least one operator in Figure 4 applies successfully at a point **p**, **p** and **n** are marked in S. After scanning of P has been completed, points in S are identified as end, link, and junction points. If a point **p** in S, which is to be identified, is a marked point, other marked neighbors of **p** are removed temporarily, and **p** is identified. Then, the removed neighboring points are restored. Let Q denote the resultant set of identified skeletal points.

The problem depicted in Figure 3(b) happens due to imperfect detection of edges. In the case of a Canny edge detector, detection of edgels sometimes fails particularly near junctions, which results in unwanted skeletal points like $\mathbf{p}_3$ in Figure 3(b). Such points cannot be removed by the method introduced in Step 4, since most of them have the value of ϕ close to $180°$. Therefore, they should be treated separately. Figure 5 illustrates a scene near the incomplete junction of edges, where $\mathbf{p}_r$ is a skeletal point to be removed, and our aim is to find $\mathbf{p}_r$. What to do first is to find a region containing $\mathbf{p}_r$, which is represented by a square R in Figure 5. The center of R is located on the end point of an edge ($\mathbf{e}_c$ in the figure), and its width w is given as twice the maximum of expected gap width between edges. A point $\mathbf{p}'_r$, which is a candidate for $\mathbf{p}_r$, is searched for in R using the assumption that $\mathbf{p}_r$ is the point closest to $\mathbf{e}_c$ among skeletal points on a line extended from $\mathbf{e}_c$ in the direction of $\mathbf{e}_l\vec{\mathbf{e}}_c$ where $\mathbf{e}_l$ is a neighboring link edgel of $\mathbf{e}_c$. Actually, $\mathbf{p}'_r$ is given by

$$\mathbf{p}'_r = \arg\min_{\mathbf{p} \in T} \|\mathbf{p} - \mathbf{e}_c\|, \quad (3)$$

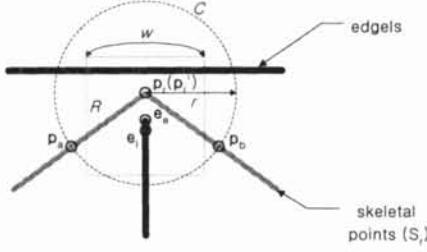Figure 4: Operators for finding skeletal points to be disconnected.

Figure 5: Scene near the incomplete junction of edges.

$$T = \{\mathbf{p}| \ \frac{\mathbf{e}_l\vec{\mathbf{e}}_e \cdot \mathbf{e}_e\vec{\mathbf{p}}}{\|\mathbf{e}_l\vec{\mathbf{e}}_e\|\|\mathbf{e}_e\vec{\mathbf{p}}\|} > \cos(\arctan \frac{2}{w}), \ \mathbf{p} \in Q_{link}, R\} \quad (4)$$

where Q_{link} represents a set of link points belonging to Q. It is an important condition in Equation 4 that the angle θ between $\mathbf{e}_l\vec{\mathbf{e}}_e$ and $\mathbf{e}_e\vec{\mathbf{p}}'_r$ should be small (i.e. $\theta < \arctan \frac{2}{w}$), since if we choose the skeletal point closest to $\mathbf{e}_e$ as $\mathbf{p}'_r$, not taking the angle into consideration, it will produce an undesirable result such as the point $\mathbf{p}_1$ in Figure 6(a).

All $\mathbf{p}'_r$'s satisfying Equation 3 are not accepted as $\mathbf{p}_r$ due to the problem illustrated in Figure 6(b). In this figure, the point $\mathbf{p}_2$ also satisfies Equation 3, but removing $\mathbf{p}_2$ seems to cause an unnatural disconnected skeleton. It is because the gap width at $\mathbf{e}_e$ is almost the same as the width of the structure represented by the skeleton. Thus, it is important to measure how rapidly the width of the structure varies in the vicinity of $\mathbf{p}'_r$. Let S_r denote a skeletal segment containing $\mathbf{p}'_r$. We can consider two points $\mathbf{p}_a$ and $\mathbf{p}_b$ ($\mathbf{p}_a, \mathbf{p}_b \in S_r$) farthest from $\mathbf{p}'_r$ within the circle C centered on $\mathbf{p}'_r$ with the radius r (see Figure 5). The points $\mathbf{p}_a$ and $\mathbf{p}_b$ can be obtained by tracking points belonging to S_r, starting from $\mathbf{p}'_r$ in opposite directions. Tracking stops when it reaches the end points of S_r or the boundary of the circle C. The last points tracked become $\mathbf{p}_a$ and $\mathbf{p}_b$. We use the parameter d in Equation 2 to measure the deviation of the structure's width within C. The radius r of C is given as $2d(\mathbf{p}'_r)$. The point $\mathbf{p}'_r$ is accepted as $\mathbf{p}_r$, if the condition

$$|d(\mathbf{p}_a) - d(\mathbf{p}'_r)| > d_f d(\mathbf{p}'_r) \ \text{ or } \ |d(\mathbf{p}_b) - d(\mathbf{p}'_r)| > d_f d(\mathbf{p}'_r) \quad (5)$$

holds where d_f is an user-specified deviation factor. Then, $\mathbf{p}_r$ is removed from Q, and its neighbors are identified again.

Step 6. pruning skeleton branches : In Step 5, some skeletal segments are split up intentionally, which may result in some noisy skeletal branches that are not useful to shape analysis. Thus, we need to prune such branches. We use Arcelli and Baja's pruning algorithm because of its simplicity and efficiency [8]. Pruning starts from an end point of each branch. For each skeletal point $\mathbf{p}$ in the branch that ends with $\mathbf{p}_e$, the quantity

$$r(\mathbf{p}, \mathbf{p}_e) = M_d(\mathbf{p}_e) - M_d(\mathbf{p}) + \|\mathbf{p} - \mathbf{p}_e\| \quad (6)$$

is computed where M_d represents the distance map obtained in Step 2. If r is less than a given threshold T_r, $\mathbf{p}$ is removed from the branch, and pruning goes on. It is stopped when either r becomes greater than or equal to T_r, or the other end of the branch is reached. The quantity $r(\mathbf{p}, \mathbf{p}_e)$ in Equation 6 can be interpreted as the loss of information we get, in terms of reconstruction of
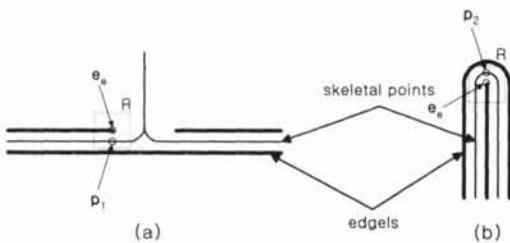

Figure 6: Two important factors in determining $\mathbf{p}_r$.

an initial shape, if the branch from $\mathbf{p}$ to $\mathbf{p}_b$ is pruned away. After pruning has been finished, very short branches dangling from main skeletal segments are deleted.

Step 7. beautifying a skeleton : Beautifying a skeleton means streightening its zigzags that are mostly caused by the unit-width thinning operation. Refer to [8] for the detailed algorithm.

3 Skeletal Segment Classification

Since we are dealing with a gray-scale image, not a binary image, it is meaningful to provide the skeleton obtained in section 2 with an ability to describe the gray-scale intensity information of underlying structures as well as their shapes. Each skeletal point can be classified as one of three types (RIDGE, RAVINE, and STAIR) according to the cross-sectional shape of the image in its neighborhood. A typical example of a cross-sectional shape for each type is shown in the first column of Figure 7. In our method, the direction and the size of the cross-section at the position of a skeletal point $\mathbf{p}$ are decided by $\mathbf{p}$ and its two nearest edgels $\mathbf{e}_p$ and $\mathbf{e}_n$ (see Figure 2). To speak more exactly, the intensity of pixels lying on two line segments $\overline{\mathbf{e}_n\mathbf{p}}$ and $\overline{\mathbf{p}\mathbf{e}_p}$ (from $\mathbf{e}_n$ to $\mathbf{e}_p$ through $\mathbf{p}$) constitute the profile of the cross-section. Let L_p denote a set of pixels on $\overline{\mathbf{e}_n\mathbf{p}}$ and $\overline{\mathbf{p}\mathbf{e}_p}$ except $\mathbf{e}_n$ and $\mathbf{e}_p$, and let $I(\mathbf{p})$ denote a gray-scale intensity value at $\mathbf{p}$. It should be noted that it is reasonable to use the Gaussian smoothed version of an input image whose value of σ is equal to that of a Canny edge detector, rather than an input gray-scale image directly, since locations of edges, which serve as boundaries of structures to be detected, are affected by the value of σ. To which type $\mathbf{p}$ belongs is determined by the following rule :

```
if (i_max(p) > e_max(p) and i_min(p) > e_min(p))
    p is labeled RIDGE;          // See Figure 7(a) and (b).
else if (i_max(p) < e_max(p) and i_min(p) < e_min(p))
    p is labeled RAVINE;         // See Figure 7(c) and (d).
else if (i_max(p) < e_max(p) and i_min(p) > e_min(p))
    p is labeled STAIR;          // See Figure 7(e) and (f).
else {
    if ((e_min(p) − i_min(p)) < (i_max(p) − e_min(p)))
        p is labeled RIDGE;
    else p is labeled RAVINE;   // See Figure 7(g).
}
```

where

$$i_{max}(\mathbf{p}) = \max_{\mathbf{q} \in L_p} I(\mathbf{q}), \ \ i_{min}(\mathbf{p}) = \min_{\mathbf{q} \in L_p} I(\mathbf{q}), \quad (7)$$

$$e_{max}(\mathbf{p}) = \max \ (I(\mathbf{e}_p), I(\mathbf{e}_n)), \ \ e_{min}(\mathbf{p}) = \min \ (I(\mathbf{e}_p), I(\mathbf{e}_n)). \quad (8)$$

The above rule can be understood intuitively by looking at Figure 7. Attention should be paid particularly to the second column of the figure where each profile represents a special case of the corrsponding type. For example, Figure 7(b) looks different from the cross-section of a typical ridge like (a), where the peak of the profile is positioned far from the center of the structure, and moreover, the intensity value at the center $\mathbf{p}$ is less than the value at the boundary point $\mathbf{e}_p$. In fact, such deviations from a standard ridge shape make a curvilinear structure detector distinguished from a simple ridge detector, as pointed out in [2].

After all skeletal points have been classified and labeled, connected points with the same label are grouped to form a skeletal segment with the single type label. Note that grouping cannot be continued beyond junction points. Skeletal segments with their length less than or equal to a threshold T_s (T_s is usually small) are relabeled UNDETERMINED because they seem to be unstable and noisy. If they are adjacent to stable segments whose length is greater than T_s, the points belonging to the unstable segments are merged iteratively to the nearest stable segments. Figure 8(a)-(d) illustrate this process. And, each isolated segment labeled UNDETERMINED is reassigned a new type such that the number of points belonging to the type is maximum in that segment. For example, a segment in Figure 8(e) gets labeled RIDGE, since points labeled RIDGE are dominant in the segment.
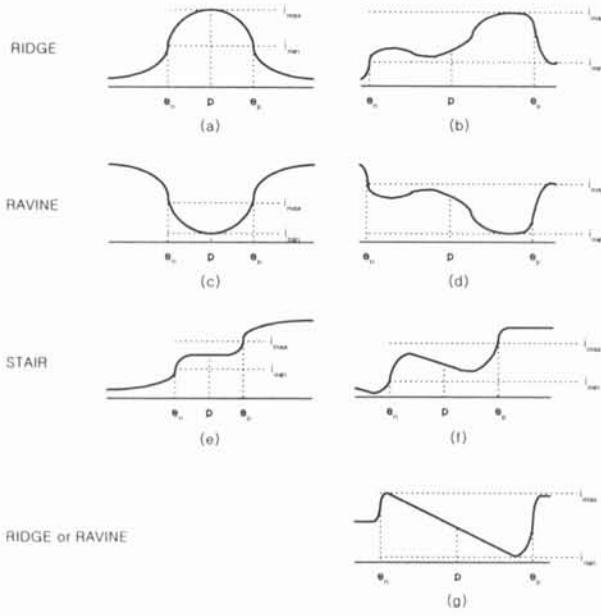
Figure 7: Examples of cross-sectional shapes for RIDGE, RAVINE, and STAIR.

4 Experimental Results and Remarks

One of the results obtained with the proposed method is shown in Figure 9, where three types of structures detectable with our method are displayed with different colors of lines. This figure illustrates main features of our method well. From the figure, one can see that our method enables us to find a variety of curvilinear structures of different width and type concurrently. Since most of conventional detectors are intended to detect curvilinear structures with uniform width such as the structure indicated by (a) in Figure 9, they have a common weakness that they detect only parts of tapering structures like (b). But, our detector is capable of finding more complex structures, e.g. (c) in the figure, as well as tapering structures.

Another result is given in Figure 10, where only ridge-like structures are marked whose width is within the range of 2 to 9 pixels. The width at a skeletal point $\mathbf{p}$ is defined as $2M_d(\mathbf{p})$. This figure shows the ability of our method to select the width of curvilinear structures to be detected.

We are now concentrating on developing an algorithm for region segmentation based on the above result. This work is closely related with the reconstruction of an initial shape from its skeleton. We are also investigating the possibility of hierarchical de-

Figure 8: Relabeling of unstable skeletal points ($T_s = 3$). (a) Construction of initial skeletal segments. (b) Segments with the length less than T_s are relabled UNDETERMINED. (c) Merging : after the first iteration. (d) Merging : after the second iteration. Merging has been finished. (e) Isolated short segment. (f) Relabeling result for the segment in (e).

☐: RIDGE   ■: STAIR   ■: RAVINE

Figure 9: An example of the results produced by the proposed algorithm. Parameters : $\sigma = 1.3, T_l = 1.0, T_h = 8.0, \phi_l = 60°, \phi_h = 120°, w = 5, d_f = 0.5, T_r = 1, T_s = 5$.

scription of an image using our method, and looking for new, feasible applications.

References

[1] Th. M. Koller, G. Gerig, G. Székely, and D. Dettwiler, "Multiscale Detection of Curvilinear Structures in 2-D and 3-D Image Data," *Proc. ICCV*, pp. 864-869, 1995.

[2] Carsten Steger, "An Unbiased Detector of Curvilinear Structures," *IEEE Trans. on PAMI*, 20(2), pp. 113-125, 1998.

[3] A. Zlotnick and P. Carnine, "Finding Road Seeds in Aerial Images," *CVGIP : Image Understanding*, 57, pp. 243-260, 1993.

[4] John Canny, "A Computational Approach to Edge Detection," *IEEE Trans. on PAMI*, 8(6), pp. 679-698, 1986.

[5] Olivier Cuisenaire, "Region Growing Euclidean Distance Transforms," *Proc. ICIAP*, Vol. 1, pp. 263-270, 1997.

[6] Carlo Arcelli and Gabriella Sanniti di Baja, "Ridge Points in Euclidean Distance Maps," *Pattern Recognition Letters*, 13, pp. 237-243, 1992.

[7] Grégoire Malandain and Sara Fernández-Vidal "Euclidean Skeletons," *Image and Vision Computing*, 16, pp. 317-327, 1998.

[8] Carlo Arcelli and Gabriella Sanniti di Baja, "Euclidean Skeleton via Centre-of-Maximal-Disc Extraction," *Image and Vision Computing*, 11(3), pp. 163-173, 1993.
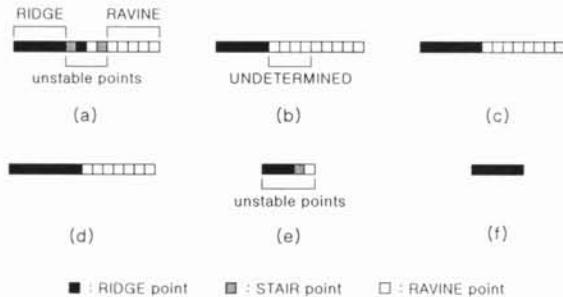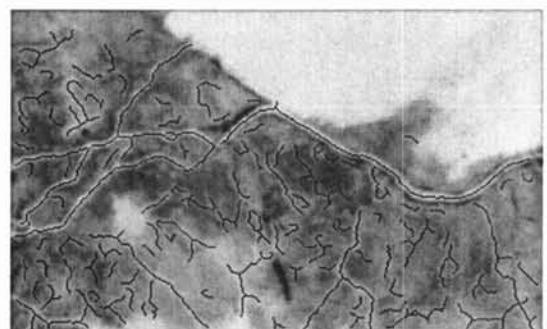
Figure 10: Another example. Parameter values used are equal to those in Figure 9.