Invited Paper

Segmentation tools in Mathematical Morphology

Serge Beucher

Centre de Morphologie Mathématique - Ecole des Mines
35 Rue Saint-Honoré, 77305 Fontainebleau (France)

ABSTRACT

This paper presents a general methodology for picture segmentation using tools provided by mathematical morphology. This methodology is based on the marking of the objects to be segmented. The marking (using techniques which may differ according to the kind of picture to be analyzed) provides a "marker set" which is used to modify the gradient of the image. This modification using geodesic image reconstruction produces a new gradient image. The main characteristic of this modified gradient is that its minima exactly fit the various connected components of the "marker set".

In a second step, a morphological transform called "watersheds" is performed on this gradient image. The watershed transform produces a partition of the image into homogeneous regions called "catchment basins". Every catchment basin contains only one marker and its boundary corresponds to the pixels of the image where the contrast is locally maximum. Thus, the transformed image exhibits the contours of the marked objects.

Some examples illustrate the use of this process when objects marking is not too complex. Then, we extend this method to situations where the marking step is not obvious, and we show how the watershed transform together with the simplification of the image can provide efficient tools for detecting homogeneous regions in an image.

INTRODUCTION

Image segmentation by mathematical morphology is performed in two steps. The first one consists in marking the objects in the image to be segmented. The second one uses a morphological tool named watershed transformation. This transformation produces a partition of the image into regions called catchment basins. The watershed transformation was applied to segmentation problems for the first time by Beucher & Lantuéjoul[1]. Unfortunately, this transformation very often leads to an over-segmentation of the image. This over-segmentation can be avoided by marking as shown by Meyer[2]. The marking of the objects provides a marker set which is used to change the homotopy of the function in the watershed transformation. This function, very often, corresponds to the gradient image, but this is not compulsory. We may use other functions depending on the problem to be solved. The first part of this paper will be devoted to the introduction of these basic tools. Some algorithms of watershed transformation will be described. This transformation will be applied to a simple case, then to a more complicated one, where over-segmentation occurs.

The concept of markers and the modification of the homotopy of the gradient function solve this problem. Other examples will be described to illustrate the methodology.

The second part of this presentation will introduce the notion of image simplification which leads to a representation of a picture in terms of graph. We will show that the basic morphological transformations can be defined on any graph, and in particular the watershed transformation. The watershed becomes then a powerful tool for a hierarchical segmentation of complex pictures.

1. BASIC TOOLS FOR SEGMENTATION IN MATHEMATICAL MORPHOLOGY

For the sake of simplicity, we will consider digital pictures only. A grey-tone image can be represented by a function $f : \mathbb{Z}^2 \rightarrow \mathbb{Z}$. $f(x)$ is the grey value of the image at point x. The points of the space $\mathbb{Z}^2$ may be the vertices of a square or an hexagonal grid.

A section of f at level i is a set $X_i(f)$ defined as :

$$X_i(f) = \{x \in \mathbb{Z}^2 : f(x) \geq i\} \qquad (1)$$

In the same way, we may define the set $Z_i(f)$:

$$Z_i(f) = \{x \in \mathbb{Z}^2 : f(x) \leq i\} \qquad (2)$$

We obviously have:

$$X_i(f) = Z^c_{i+1}(f) \qquad (3)$$

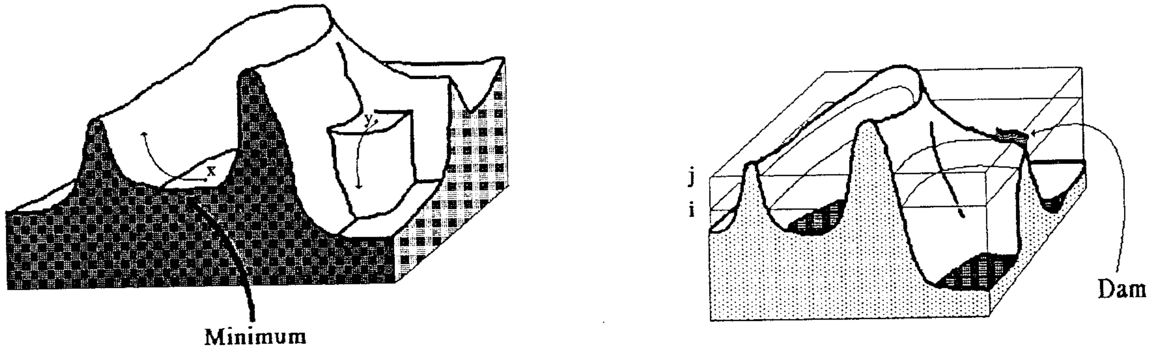
1.1. The watershed transformation

The set of all the points $\{x, f(x)\}$ belonging to $\mathbb{Z}^2 \times \mathbb{Z}$ can be seen as a topographic surface S. The lighter the grey value of f at point x, the higher the altitude of the corresponding point $\{x, f(x)\}$ on the surface.

Various characteristic features can be defined on this topographic surface. Among them are the **minima** also called **regional minima** of f. Consider two points s_1 and s_2 of this surface S. A path between $s_1(x_1, f(x_1))$ and $s_2(x_2, f(x_2))$ is any sequence $\{s_i\}$ of points of S, with s_i adjacent to s_{i+1}. A non ascending path is a path where :

$$\forall\ s_i(x_i, f(x_i)),\ s_j(x_j, f(x_j)) \quad i \geq j \ \Leftrightarrow\ f(x_i) \leq f(x_j) \qquad (4)$$

A point $s \in S$ belongs to a minimum iff there exists no ascending path starting from s. A minimum can be considered as a sink of the topographic surface (Fig.1a). The set M of all the minima of f is made of various connected components $M_i(f)$. Let us define the **catchment basins** of f and the **watershed lines** by means of a flooding process. Imagine that we pierce each



(a) (b)

Fig.1.Minima of a function (a), flooding of the relief and construction of dams (b).

minimum $M_i(f)$ of the topographic surface S, and that we plunge this surface into a lake with a constant vertical speed. The water entering through the holes floods the surface S. During the flooding, two or more floods coming from different minima may merge. We want to avoid this phenomenon and we build a dam on the points of the surface S where the floods would merge (Fig.1b). At the end of the process, only the dams emerge. These dams define the watershed of the function f. They separate the various catchment basins $CB_i(f)$, each one containing one and only one minimum $M_i(f)$ (Fig.2).

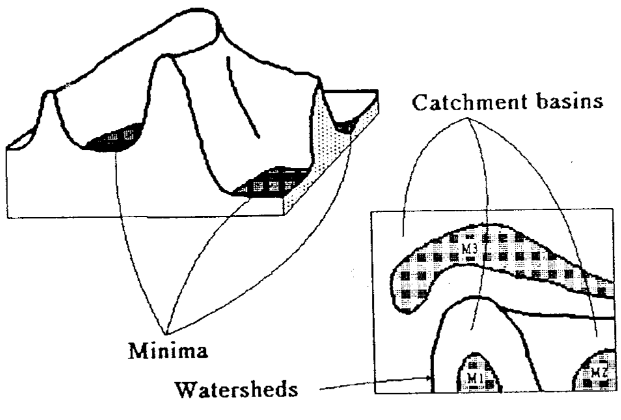


Fig.2.Regional minima, catchment basins and divide lines.

1.2. Building the watershed

This definition of the watershed transformation has a great advantage : it can be performed on the sections of the function. The watershed algorithms can be divided into two groups. The first group contains algorithms which simulate the flooding process. The second group consists of procedures aiming at detecting directly the points of the watershed lines.

Let us describe the algorithm based on the sections (it belongs to the first group). But, before, we must introduce some geodesic transformations.

1.2.1. Geodesy, geodesic zones of influence

Let $X \subset \mathbb{Z}^2$ be a set, and x and y be two points of X. We define the **geodesic distance** $d_X(x,y)$ between x and y as the length of the shortest path (if it exists) included in X and linking x and y (Fig.3a).

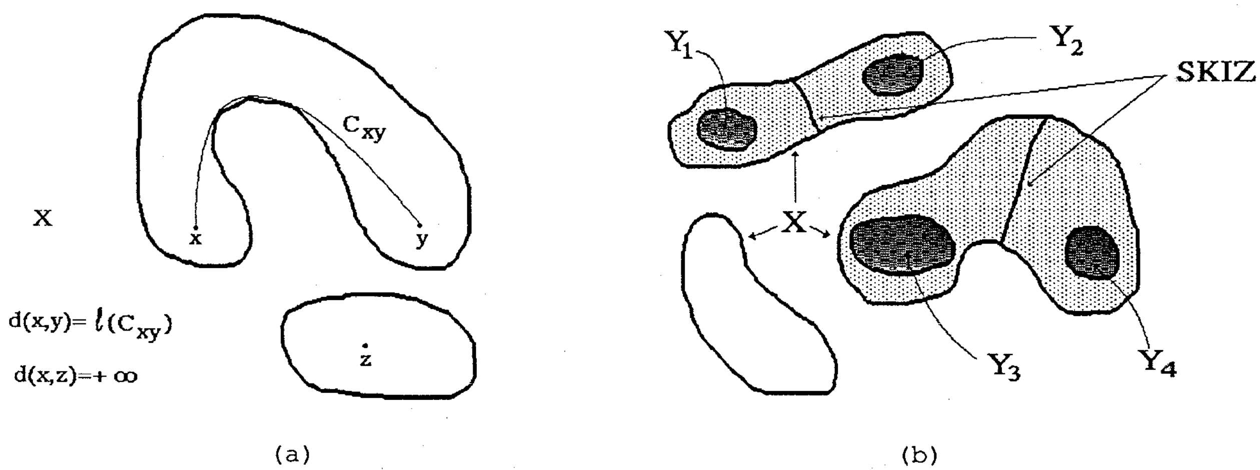


Fig.3. Shortest path and geodesic distance (a), Geodesic skeleton by zones of influence (b).

Many basic morphological transforms can be redefined with this distance. Let Y be any set included in X. We can compute the set of all points of X which are at a finite geodesic distance from Y :

$$R_X(Y) = \{x \in X : \exists\, y \in Y,\ d_X(x,y) \text{ finite}\} \qquad (5)$$

$R_X(Y)$ is called the **X-reconstructed set** by the marker set Y. It is made of all the connected components of X which are marked by Y.

Suppose now that Y is composed of n connected components Y_i. The **geodesic zone of influence** $z_X(Y_i)$ of Y_i is the set of the points of X that are at a finite geodesic distance from Y_i and closer to Y_i than to any other Y_j (Fig.3b) :

$$z_X(Y_i) = \{x \in X : d_X(x,Y_i) \text{ finite and } \forall j \neq i,\ d_X(x,Y_i) < d_X(x,Y_j)\} \qquad (6)$$

The boundaries between the various zones of influence give the **geodesic skeleton by zones of influence** of Y in X.

We shall write :

$$SKIZ_X(Y) = \bigcup_i z_X(Y_i) \qquad (7)$$

1.2.2. Watersheds and sections of f

Consider (Fig.4) a section $Z_i(f)$ of f at level i, and suppose that the flood has reached this height. Consider now the section $Z_{i+1}(f)$. We immediately see that the flooding of $Z_{i+1}(f)$ is performed in the zones of influence of the connected components of $Z_i(f)$ in $Z_{i+1}(f)$. Some connected components of $Z_{i+1}(f)$ which are not reached by the flood are, by definition, minima at level i+1. These minima must therefore be added to the flooded area. Denoting by $W_i(f)$ the section at level i of the catchment basins of f, and by $M_{i+1}(f)$ the minima of the function at height i+1, we have :

$$W_{i+1}(f) = [\, SKIZ_{Z_{i+1}(f)}(X_i(f))\,] \cup M_{i+1}(f) \qquad (8)$$

The minima at level i+1 are given by :

$$M_{i+1}(f) = Z_{i+1}(f) \,/\, R_{Z_{i+1}(f)}(Z_i(f)) \qquad (9)$$

where / stands for the set difference.

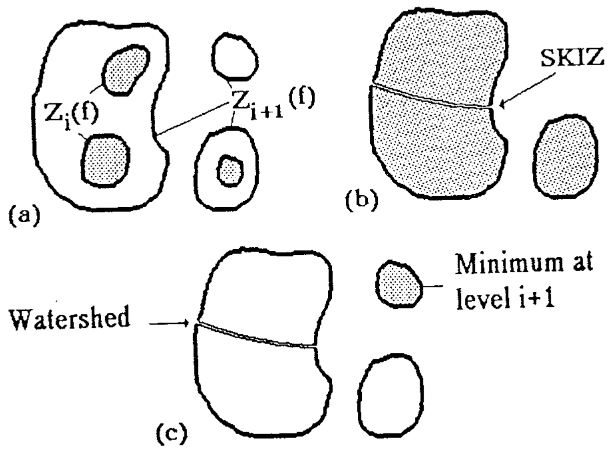


Fig.4.Successive steps of the watershed construction using geodesic SKIZ.

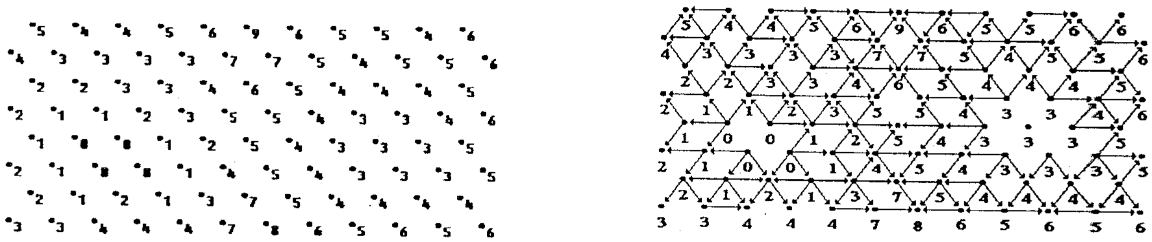
This iterative algorithm is initialized with $W_{-1}(f) = \emptyset$. At the end of the process, the watershed line DL(f) is equal to :

$$DL(f) = W_N^c(f) \quad (\text{with } \max(f) = N) \qquad (10)$$

1.2.3. Arrowing and watersheds

The previous algorithm belongs to the first group : it simulates the flooding of the surface S starting from the minima of f. We will now present briefly another algorithm that belongs to the second group and which is based on the **arrowing representation** of a function f.

From $f_2 : \mathbb{Z}^2 \rightarrow \mathbb{Z}$, we may define an oriented graph whose vertices are the points of $\mathbb{Z}^2$ and equipped with edges (or arrows) from x to any adjacent point y iff $f(x) < f(y)$ (Fig.5)[3].



function f

Complete graph of arrows

(a) (b)

Fig.5.Function f (a) and its complete graph of arrows (b).

The definition does not allow the arrowing of the plateaus of the topographic surface. This arrowing can be performed by means of geodesic dilations. The operation is called the **completion** of the graph of arrows. We may, then, select on the complete graph some configurations which, locally, correspond to divide lines. These configurations are represented on Fig.6 for the 6-connectivity neighborhood of a point (up to a rotation).

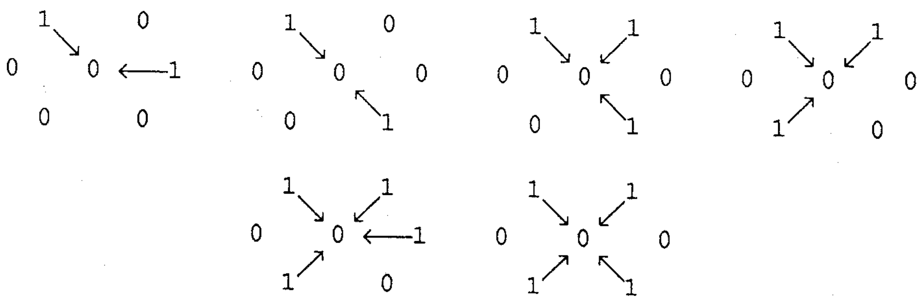
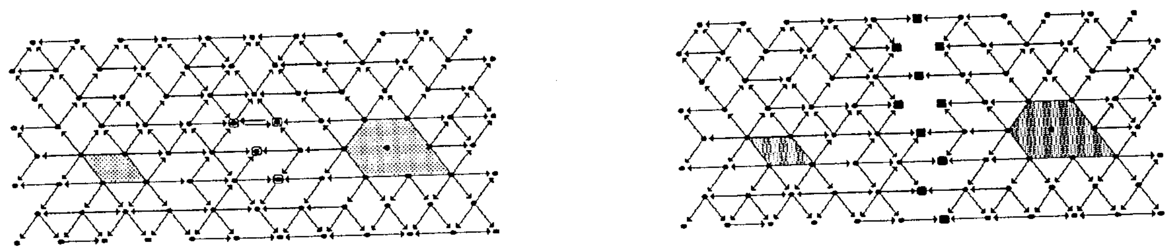


Fig.6.Configurations of arrows corresponding to possible divide points.

Any point receiving arrows from more than one connected component of its neighborhood may be flooded by different lakes. Consequently, this point may belong to a divide line. In a second step, the arrows starting from the selected points must be suppressed. These points, in fact, cannot be flooded, so they cannot propagate the flood. Doing so, we change the arrowing of the neighbor points. Some new divide points may then appear. The procedure is re-run until no new divide point is selected (Fig.7).

This algorithm produces local watershed lines. The true watersheds can be extracted easily. They are the only ones that form closed curves.



Selection of primary points

(a)

Final result

(b)

Fig.7.Watersheds by arrowing. (a) Selection of primary divide points, (b) final result and corresponding graph of arrows.

2. USE OF WATERSHEDS IN PICTURE SEGMENTATION

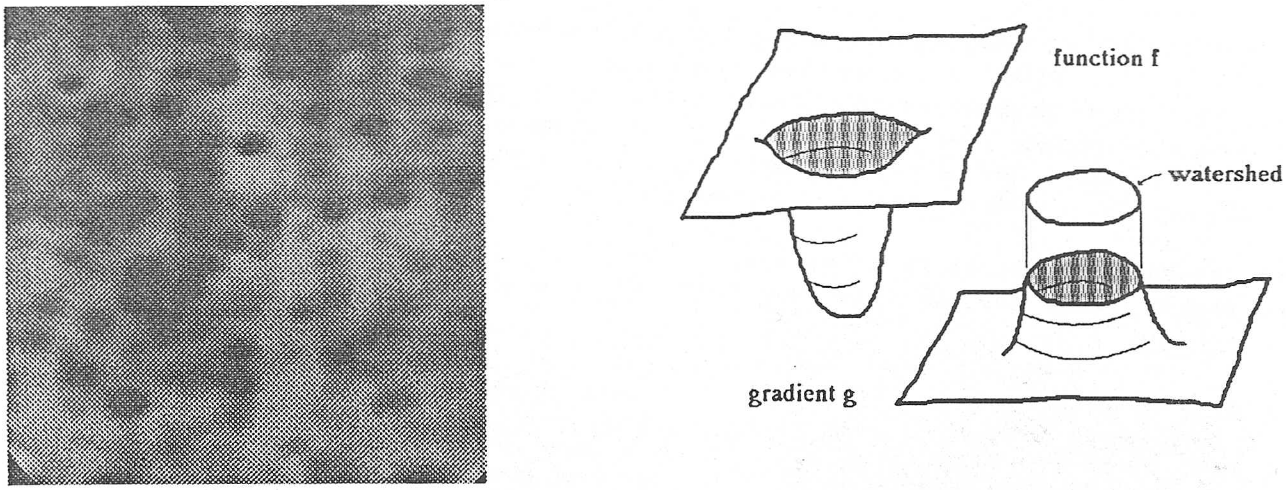
2.1. A real example

The use of watershed transformations for picture segmentation will be explained by means of a real example : the contouring of proteins in an electrophoresis gel (Fig.8a). Each blob can be characterized by a sink in the topographic surface drawn by the grey-tone image f. The corresponding gradient image should present a volcano-type topography as depicted in Fig.8b. The contours of the proteins blobs correspond therefore to the watershed lines of the gradient image g(f). This gradient is defined as :

$$g(f) = (f \oplus B) - (f \ominus B) \qquad (11)$$

where $f \oplus B$ and $f \ominus B$ are respectively elementary dilation and erosion of f[4].

Unfortunately, the real watershed transform of the gradient, given in Fig.9a, presents many catchment basins; there are as many of them as there are minima in the gradient. These minima are produced by small variations in the grey values. This over-segmentation could be reduced by appropriate filtering. But a better result will be obtained if we mark the patterns to be segmented



(a) (b)

Fig.8.Electrophoresis gel (a), topographic surface of the initial function and of the gradient image (b).

before performing the watershed transformation of the gradient. Suppose that we mark each blob of protein of the Figure 8a. This marking can be performed by extracting the minima of f. We must also define a marker for the background. In order to get a connected marker surrounding the blobs, the watershed transform of the initial image is performed. Then, we obtain a set of markers M (Fig.9b). We consider again the topographic surface of the gradient image and the flooding process, but, instead of piercing the minima of this surface, we will only make holes through the components of the marker set M. The flooding will invade the surface and produce as many catchment basins as there are markers comprised in the markers set. Moreover, the watershed lines will correspond to the crest lines of this topographic surface which themselves correspond to the contours of the objects (Fig.9c).

We can show that this algorithm may be written as follows.

If $W_i(g)$ is the section at level i of the new catchment basins of g, we have :

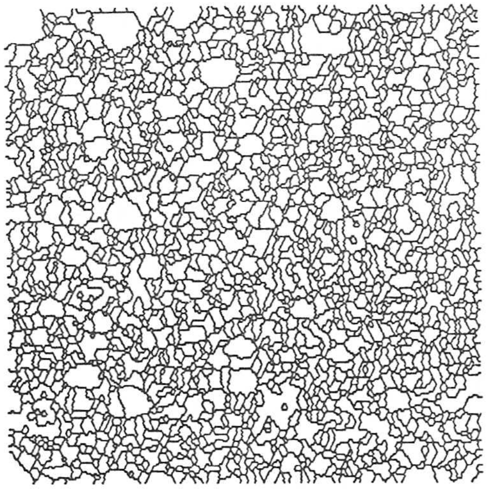
$$W_{i+1}(g) = SKIZ_{Z_{i+1} \cup M}(W_i(g)) \qquad (12)$$

with :

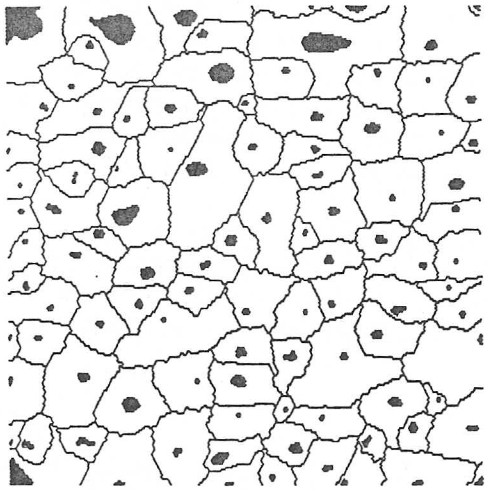
$$W_{-1}(g) = M, \text{ markers set}$$

Surprisingly, this algorithm is simpler than the pure watershed algorithm, because we do not take the real minima of g into account.

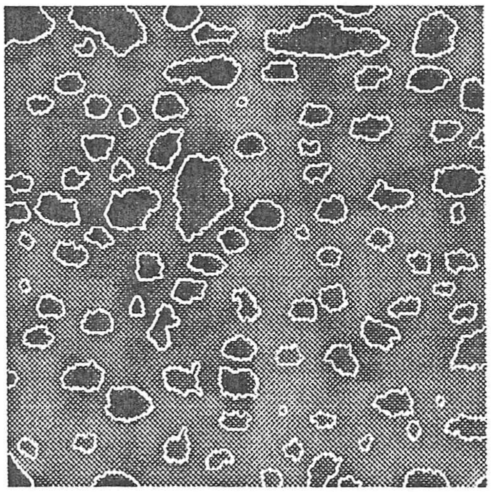
This procedure may be split into two steps. The first one consists in modifying the gradient function g in order to produce a new gradient g'. This new image is very similar to the original one, except that its initial minima have disappeared and have been replaced by the set M. This image modification also called **homotopy modification** can be performed by reconstructing the sections of g with the markers M.



(a)



(b)



(c)

Fig.9.Watersheds of the gradient image (a), set of selected markers M (inner markers of blobs, and one outer marker for the background) (b), final segmentation (c).

We have :

$$\forall i, \quad Z_i(g') = R_{Z_i(g) \cup M}(M) \qquad (13)$$

The second step simply consists in performing the watershed transform of the modified gradient g'[2].

2.2. Towards a methodology of the segmentation

This first example of segmentation leads to a general scheme. Image segmentation consists in selecting first a marker set M pointing out the objects to be extracted, then a function f quantifying a segmentation criterion (this criterion can be, for instance, the changes in grey values). This function is modified to produce a new function f′ having as minima the set of markers M. The segmentation of the initial image is performed by the watershed transform of f′ (Fig.10).

The segmentation process is therefore divided into two steps : an "intelligent" part whose purpose is the determination of M and f, and a "straightforward" part consisting in the use of the basic morphological tools which are watersheds and image modification.

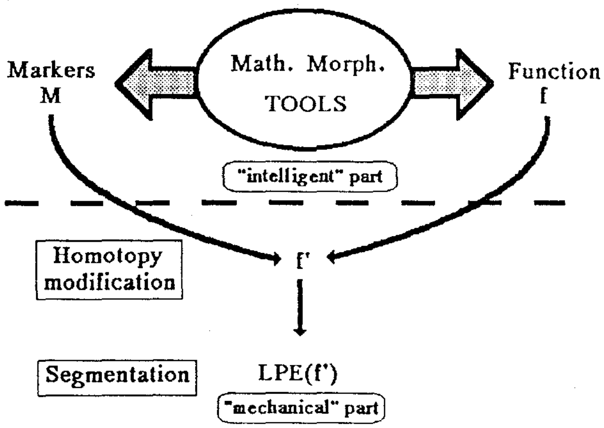


Fig.10.Synopsis of the morphological segmentation methodology.

A lot of segmentation problems may be solved according to this general scheme[5]. Let us illustrate this procedure with two examples.

The first example is a part of a road traffic sensor able to compute various parameters from a road traffic scene for each traffic lane[6]. To achieve the segmentation of the traffic lanes, two images are generated. The first one is an average picture of a sequence of images (Fig.11a) and the second one is the average of the first derivative of the sequence along the time axis (Fig.11b). These two images are used to produce a marker set (Fig.11c) and the function to be segmented (Fig.11d). The latter function is the geodesic distance function of the ground layout. The final segmentation is given in Fig.11e.

The second example is a problem of segmentation of cleavage facets in a SEM micrograph of a steel fracture (Fig.12a). The function used for watersheds (Fig.12b) along with the marker set (Fig.12c) are built by combining a photometric criterion (contrast between facets due to blazing ridges) and a shape criterion (facets are supposed to be more or less convex). The final

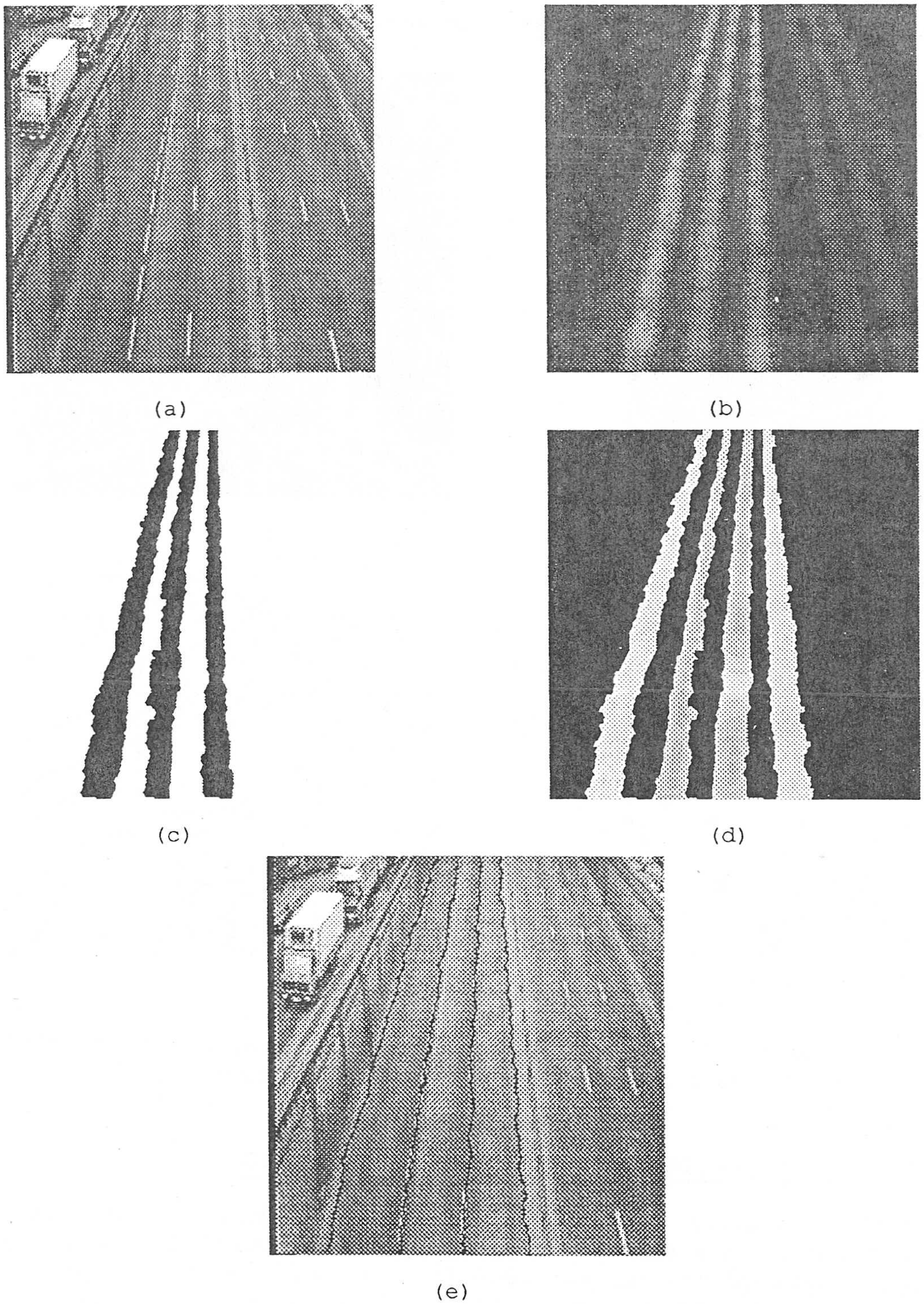
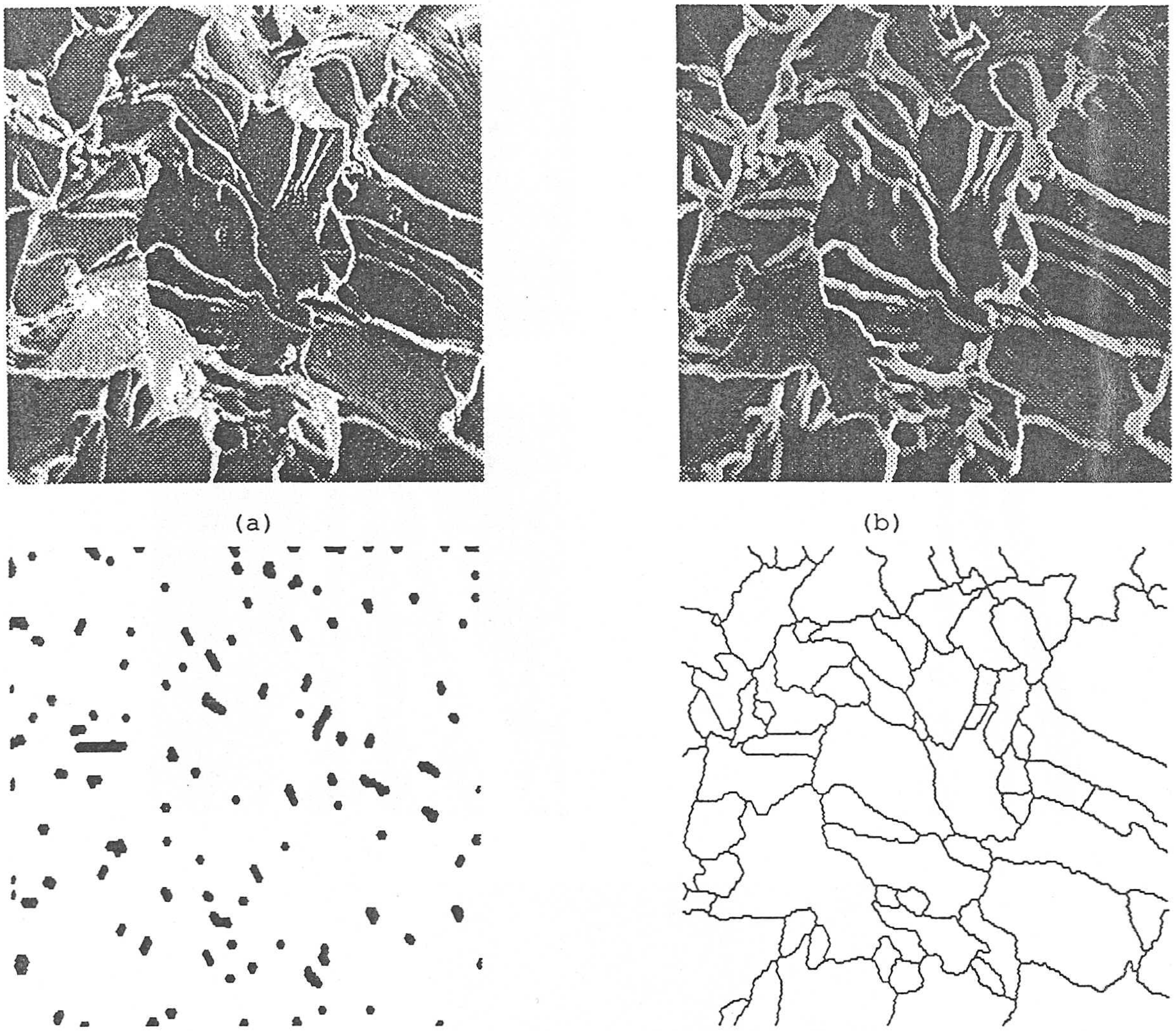


Fig.11.First image (a), second image (b), markers of traffic lanes (c), geodesic distance of ground layout (d), segmentation of the traffic lanes (e).

result (Fig.12d) is obtained after a watershed transformation and an elimination of some irrelevant arcs separating markers belonging to the same facet[5].



(a) (b) (c) (d)

Fig.12.SEM image of a metallic fracture (a), function used for segmentation (b), markers of facets (c), final result with facets detection (d).

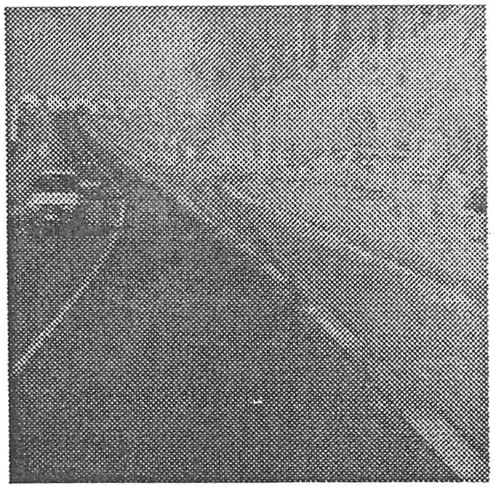
3. WATERSHEDS AND HIERARCHICAL SEGMENTATION

Very often, especially for complex and noisy pictures, the markers selection is difficult. To overcome this problem, we may simplify the initial image and try to extract homogeneous regions from this simplified picture. Both image simplification and region extraction make an intensive use of the watershed transformations.

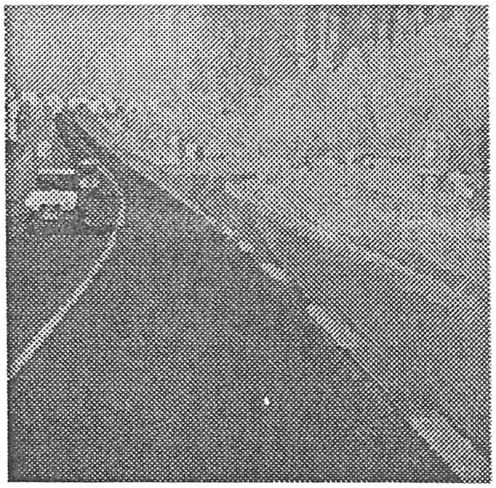
3.1. Image simplification

Consider a grey-tone image f, and its corresponding morphological gradient image g(f). A simplified image can be computed in the following way :

- First, we calculate the watersheds of the gradient image.
- Second, we label every catchment basin of the watershed transform with the grey value in the initial image f corresponding to the minima of g(f).



(a)



(b)

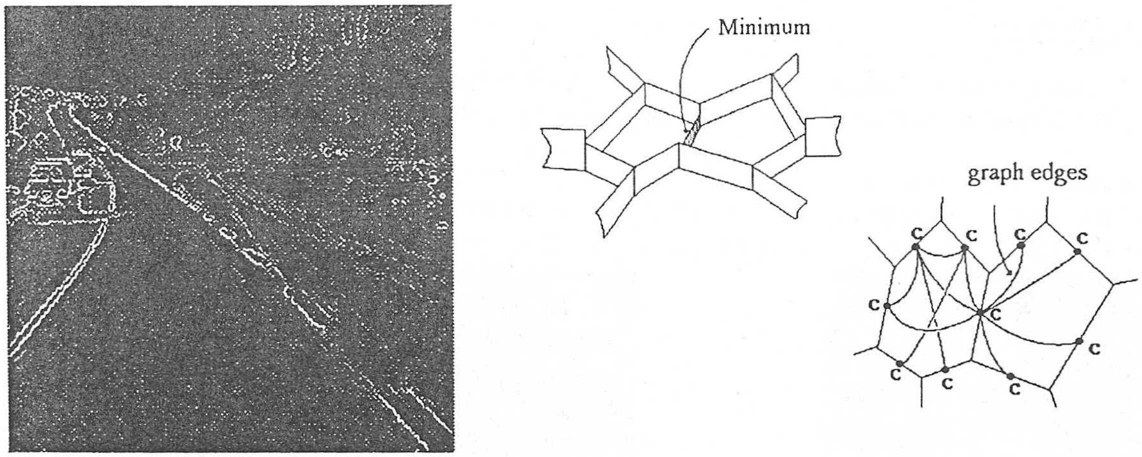
Fig.13.Original image (a), simplified mosaic image (b).

The result is a simplified image (Fig.13), made of a mosaic of pieces (the catchment basins) of constant grey levels, where no information regarding the contours has been lost. Then, this simplified image, also called **mosaic image**, may be used to define a valued graph, on which the morphological transformations, and in particular the watersheds, can be extended[7].

3.2. Hierarchical segmentation

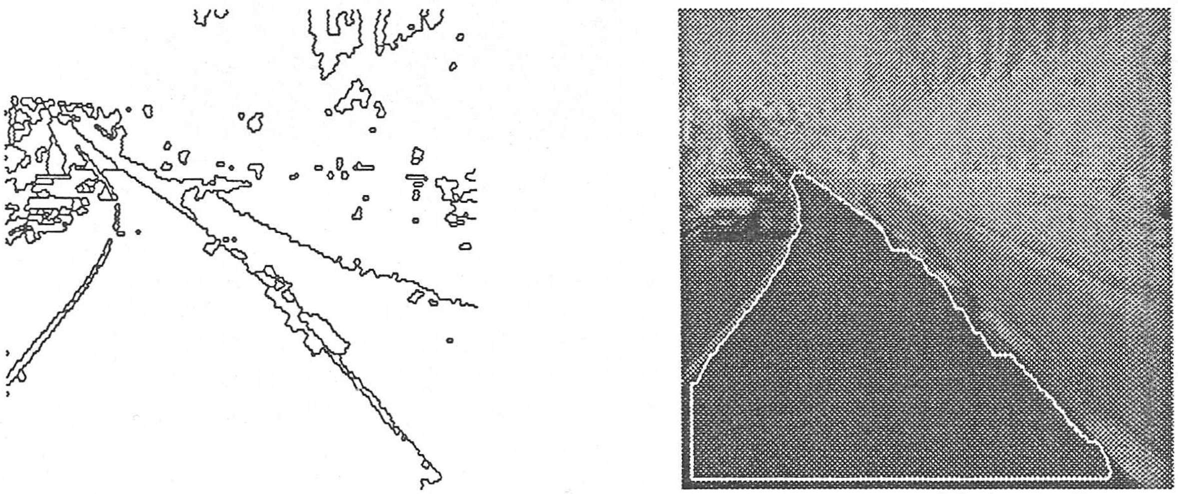
Consider the previous mosaic image. First, the boundaries between adjacent tiles of the mosaic image are valued with the difference of their grey values. This produces a gradient image whose support is the watersheds of the image g(f). Each wall of this gradient image is considered as a vertex of a valued graph (Fig.14b). Two vertices of this graph are neighbors if the corresponding boundaries surround the same tile of the mosaic picture. A watershed transformation can be applied to this graph. The weakest boundaries of the mosaic image correspond to regional minima of our valued graph (Fig.14a).

The result of the watershed transformation leads to a hierarchical segmentation of the image, as illustrated with the example given in Figure 15. A selection of markers can be made at this level to segment features in the image (for example, the road in our case). Further levels of hierarchy may also be defined by iterating this procedure[8].



(a) (b)

Fig.14.Gradient of the mosaic image (a), corresponding graph used for the hierarchical segmentation (b).



(a) (b)

Fig.15.Detection of the road : (a) First level of hierarchy, (b) marker of the road.

CONCLUSION

The watershed transformation can be considered as an all-purpose tool in the approach of segmentation problems by mathematical morphology. This tool, combined with the marker selection of the features to be extracted, leads to a general methodology for the segmentation process. First, by means of the markers selection, we point out what we want to extract from the picture. Then, we have to define the criteria used to segment the image. This means that image segmentation cannot be performed accurately and adequately if we do not build the objects we want to detect. In this approach, the picture segmentation is not the primary step of image understanding. On the contrary, a fair segmentation can be obtained only if we know exactly what we are looking for in the image.

REFERENCES

1.S.Beucher and C.Lantuéjoul, "Use of watersheds in contour detection", Int. Workshop on Image Processing, Real-Time edge and motion detection/estimation, CCETT/INSA/IRISA, IRISA Report n°132, pp. 2.1-2.12, Rennes, France, 17-21 Sept. 1979.

2.F.Meyer and S.Beucher, "Morphological Segmentation", to be published in Journal of Visual Communication and Image representation.

3.F.Maisonneuve, "Sur le partage des eaux", CMM Internal report, Paris School of Mines, Dec. 1982.

4.S. Beucher and F.Meyer, "Méthodes d'analyse des contrastes à l'analyseur de texture", Proc. of Congrès AFCET/IRIA, Reconnaissance des Formes et Traitement des Images, Tome 1, pp. 378-384, Chatenay-Malabry, France, 21-23 Feb. 1978.

5.S.Beucher, *Segmentation d'images et morphologie mathématique*, Doctorate Thesis, Paris School of Mines, France, June 1990.

6.S.Beucher, J.M.Blosseville and F.Lenoir, "Traffic spatial measurements using video image processing", Proc. of SPIE, Advances in intelligent robotics systems, Cambridge symposium on Optimal and Optoelectronic Engineering, Vol. 848, pp. 648-655, Cambridge, Mass., U.S.A., 1-6 Nov. 1987.

7.L.Vincent, "Graphs and Mathematical Morphology", Signal Processing, Vol. 16, n° 4, pp. 365-388, April 1989.

8.S.Beucher, M.Bilodeau and X.Yu, "Road Segmentation by watershed algorithms", Proc. of PROMETHEUS workshop, Sophia-Antipolis, France, April 1990 (in press).