Scanning Microscopy

Volume 1992
Number 6 *Signal and Image Processing in Microscopy and Microanalysis*

Article 28

1992

The Watershed Transformation Applied to Image Segmentation

Serge Beucher
Ecole Nationale Supérieure des Mines de Paris, France, beucher@cmm.ensmp.fr

Scanning Microscopy Supplement 6, 1992 (pages 299-314)
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

0892-953X/92$5.00+.00

THE WATERSHED TRANSFORMATION APPLIED TO IMAGE SEGMENTATION

Serge Beucher

Centre de Morphologie Mathématique, Ecole Nationale Supérieure des Mines de Paris,
35, rue Saint-Honoré, 77305 FONTAINEBLEAU CEDEX, France
Telephone Number: 33-1-64694706; FAX: 64694707; e-mail: beucher@cmm.ensmp.fr

Abstract

Image segmentation by mathematical morphology is a methodology based upon the notions of watershed and homotopy modification.This paper aims at introducing this methodology through various examples of segmentation in materials sciences, electron microscopy and scene analysis.

First, we defined our basic tool, the watershed transform. We showed that this transformation can be built by implementing a flooding process on a greytone image. This flooding process can be performed by using elementary morphological operations such as geodesic skeleton and reconstruction. Other algorithms are also briefly presented (arrows representation).

Then, the use of this transformation for image segmentation purposes is discussed. The application of the watershed transform to gradient images and the problems raised by over-segmentation are emphasized. This leads, into the third part, to the introduction of a general methodology for segmentation, based on the definition of markers and on a transformation called homotopy modification. This complex tool is defined in detail and various types of implementations are given.

Many examples of segmentation are presented. These examples are taken from various fields: transmission electron microscopy, scanning electron microscopy (SEM), 3D holographic pictures, radiography, non destructive control and so on.

The final part of this paper is devoted to the use of the watershed transformation for hierarchical segmentation. This tool is particularly efficient for defining different levels of segmentation starting from a graph representation of the images based on the mosaic image transform. This approach will be explained by means of examples in industrial vision and scene analysis.

Key Words: watershed transformation, markers, homotopy, distance function, morphological gradient, hierarchical segmentation, mosaic image, geodesy, catchment basins, arrows representation.

Introduction

The *watershed transformation* is a powerful tool for image segmentation. In this paper, the different morphological tools used in segmentation are reviewed, together with an abundant illustration of the methodology through examples of image segmentation coming from various areas of image analysis.

There exist two basic ways of approaching image segmentation. The first way is boundary-based and detects local changes. The second is region-based and searches for pixel and regional similarities. We shall see that the watershed transformation belongs to the latter class.

Beucher and Lantuejoul were the first to apply the concept of watershed and divide lines to segmentation problems [3]. They used it to segment images of bubbles and SEM metallographic pictures.

Unfortunately, this transformation very often leads to an over-segmentation of the image. To overcome this problem, a strategy has been proposed by Meyer and Beucher [7]. This strategy is called marker controlled segmentation.

This approach is based on the idea that machine vision systems often roughly "know" from other sources the location of the objects to be segmented. This approach is applied as follows: first, we define the properties which will be used to mark the objects. These markers are called *object markers*. The same is done for the background, i.e., for portions of the image in which we are sure there is no pixel belonging to any object. These markers constitute the *background markers*. The rest of the procedure is straightforward and is the same for all applications: the gradient image is modified in order to keep only the most significant contours in the areas of interest between the markers. This gradient modification consists in changing the *homotopy* of the function. Then, we perform the final contour search on the modified gradient image by using the watershed transformation. No supervision, no parameter and no heuristics are needed to perform the final segmentation. The parameterization controlling the segmentation is concentrated in the marker construction step where it is easier to control and validate it.

The gradient image is often used in the watershed transformation, because the main criterion of the segmentation is the homogeneity of the grey values of the objects present in the image. But, when other criteria are relevant, other functions can be used. In particular, when the segmentation is

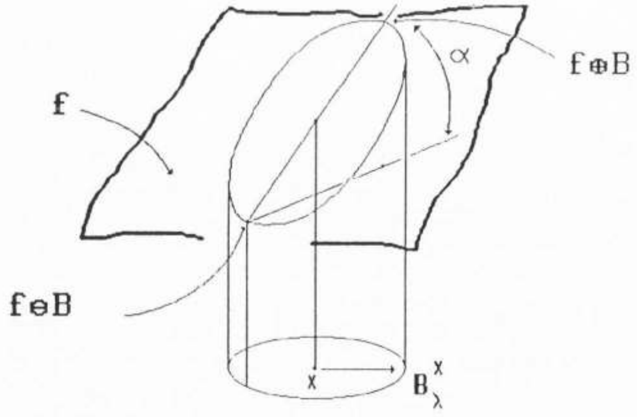


Figure 1. Construction of the morphological gradient of an image.

based on the shape of the objects, the distance function is very helpful.

In the first part, we described the main morphological tools used in segmentation: gradient, distance function, geodesic distance function and watershed transformation. For this last transformation, some algorithms are presented.

In the second part, we introduced the concept of markers and the homotopy modification of the transformed function for solving over-segmentation problems. Many examples illustrate this methodology.

The final part of this paper is devoted to the use of the watershed transformation for *hierarchical segmentation*. This tool is particularly efficient for defining different levels of segmentation starting from a graph representation of the images based on the *mosaic image transform*. This approach will be explained by means of examples in industrial vision and scene analysis.

The Basic Tools for Segmentation

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

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

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

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

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

We have obviously:

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

Morphological gradient

The *morphological gradient* of a picture is defined as:

$$g(f) = (f \oplus B) - (f \ominus B) \tag{4}$$

where $f \oplus B$ and $f \ominus B$ are respectively the elementary *dilation* and *erosion* of f by the smallest regular structuring element B defined on the digitization grid (elementary hexagon or square) [1].

When f is continuously differentiable, this gradient is equivalent to the modulus of the gradient of f (figure 1):

$$g(f) = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 \right]^{1/2} \tag{5}$$

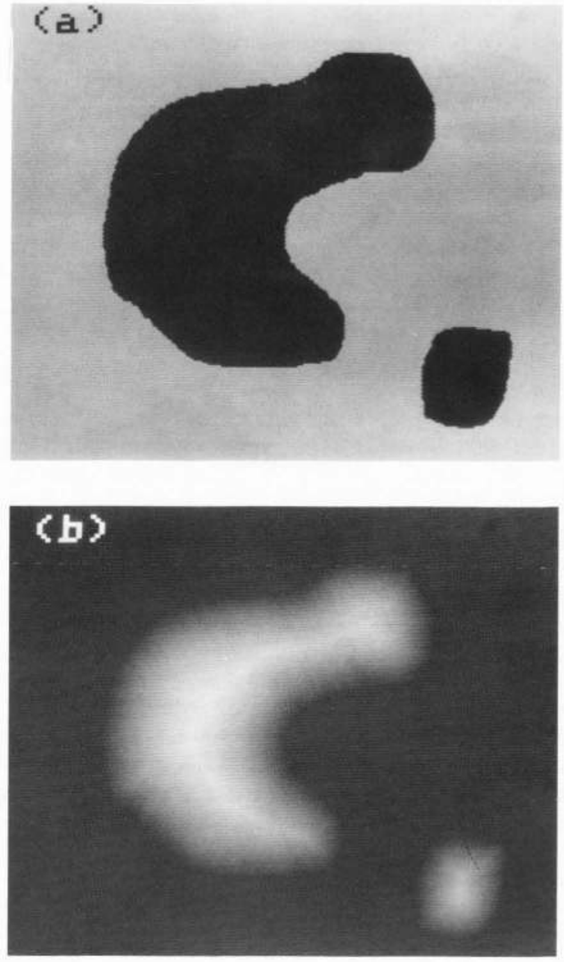


Figure 2. Distance function (b) of a set (a).

The simplest way to approximate this modulus is to assign to each point x the difference between the highest and the lowest pixels within a given neighborhood of x. In other words, for a function f, it is the difference between the dilated function $f \oplus B$ and the eroded function $f \ominus B$.

Distance function

Let Y be a set of Z^2. For every point y of Y, define the distance of y to the complementary set Y^c (figure 2):

$$\forall y \in Y,\ d(y) = \mathrm{dist}(y, Y^c) \tag{6}$$

It can easily be shown that a section of d at level i is given by:

$$X_i(d) = \{y : d(y) \geq i\} = Y \ominus B_i \tag{7}$$

where B_i is a disk of radius i.

Geodesy, geodesic distance

Let $X \subset Z^2$ be a set, x and y 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 any) included in X and linking x and y (figure 3a) [4].

Let Y be any set included in X. We can compute the set of all points of X that are at a finite geodesic distance from Y:

$$R_X(Y) = \{x \in X : \exists\, y \in Y,\ d_X(x,y) \text{ finite}\} \tag{8}$$

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

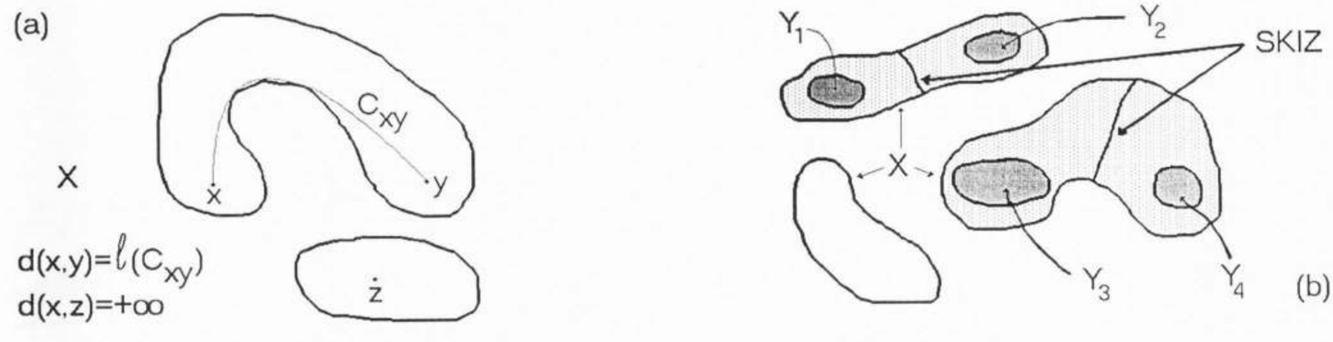


Figure 3. Shortest path and geodesic distance (a), SKIZ of a set Y in X (b).

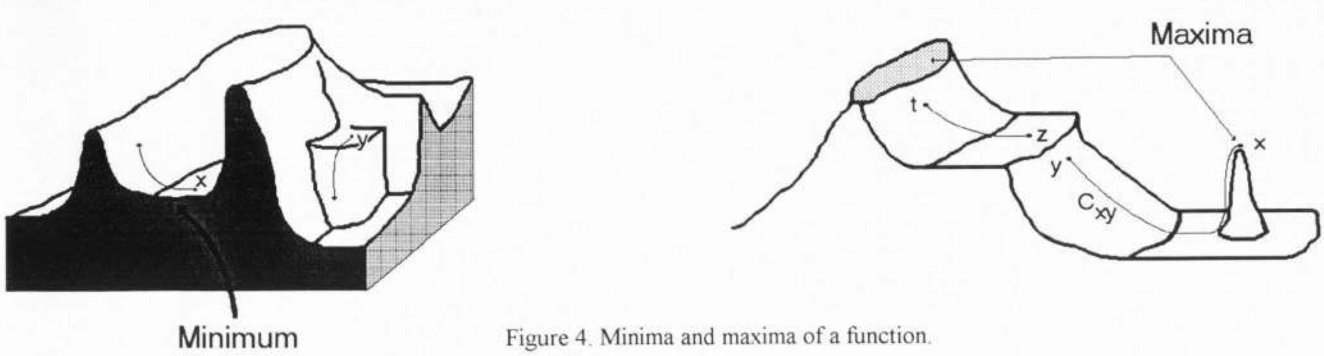


Figure 4. Minima and maxima of a function.

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 is the set of points of X at a finite geodesic distance from Y_i and closer to Y_i than to any other Y_j (figure 3b):

$$z_X(Y_i)=\{x\in X: d_X(x,Y_i) \text{ finite } \& \ \forall j\neq i,\ d_X(x,Y_i)<d_X(x,Y_j)\} \quad (9)$$

The boundaries between the various zones of influence give the geodesic skeleton by zones of influence $SKIZ_X$ of Y in X. We shall write:

$$IZ_X(Y) = \bigcup_i z_X(Y_i) \qquad (10)$$

and:

$$SKIZ_X(Y) = X / IZ_X(Y) \qquad (11)$$

where / stands for the set difference.

Minima, maxima of a function

Among the various features that can be extracted from an image, the *minima* and the *maxima* are of primary importance.

The set of all the points $\{x,f(x)\}$ belonging to $Z^2 \times Z$ 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.

The minima of f, also called *regional minima*, are defined as follows.

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:

$$\forall\ s_i(x_i, f(x_i)), s_j(x_j, f(x_j)) \quad i \geq j \Leftrightarrow f(x_i) \leq f(x_j) \qquad (12)$$

A point $s \in S$ belongs to a minimum if and only if there exist no ascending path starting from s. A minimum can be considered as a sink of the topographic surface (figure 4). The set M of all the minima of f is made of various connected components $M_i(f)$.

A similar definition holds for the maxima.

The watershed transformation

Consider again an image f as a topographic surface and define the *catchment basins* of f and the *watershed lines* by means of a flooding process. Imagine that we pierce each 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 event and we build a dam on the points of the surface S where the floods would merge. 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 $M_i(f)$ (figure 5).

Building the watershed

The definition of the watershed transformation by flooding may be directly transposed by using the sections of the function f.

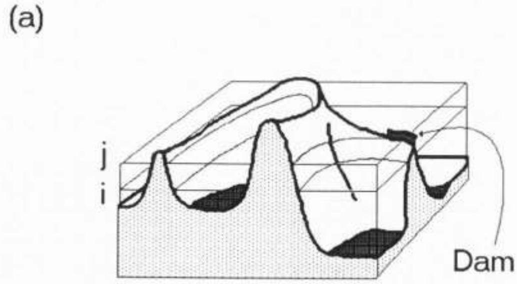
Consider (figure 6) a section $Z_i(f)$ of f at level i, and suppose that the flood has reached this height. Consider now the section $Z_{i+1}(f)$. We see immediately that the flooding of $Z_{i+1}(f)$ is performed in the zones of influence of the connected components of $Z_i(f)$ in $Z_{i+1}(f)$. Some 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:

$$W_{i+1}(f) = [IZ_{Z_{i+1}(f)}(X_i(f))] \cup M_{i+1}(f) \qquad (13)$$

The minima at level i+1 are given by:

$$M_{i+1}(f) = Z_{i+1}(f) / R_{Z_{i+1}(f)}(Z_i(f)) \qquad (14)$$

This iterative algorithm is initiated with $W_{-1}(f) = \varnothing$. At



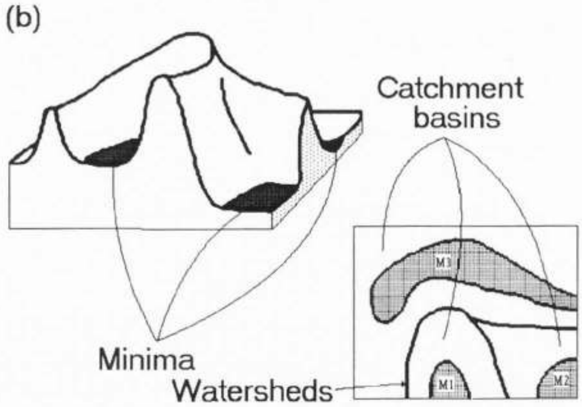


Figure 5. Flooding of the relief and dam building (a), catchment basins and divide lines (b).

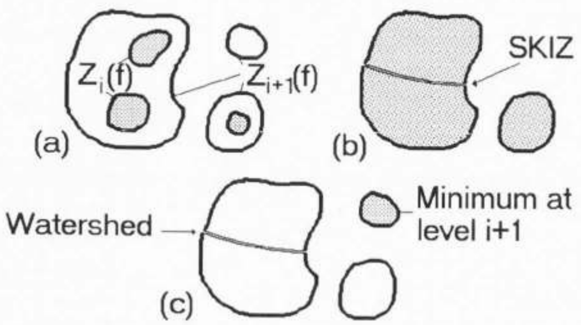


Figure 6. Watershed construction using geodesic SKIZ.

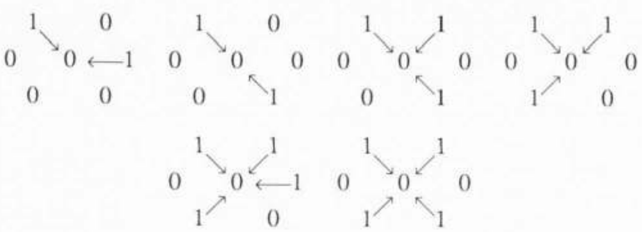


Figure 8. Configurations of arrows corresponding to possible divide points (hexagonal grid).

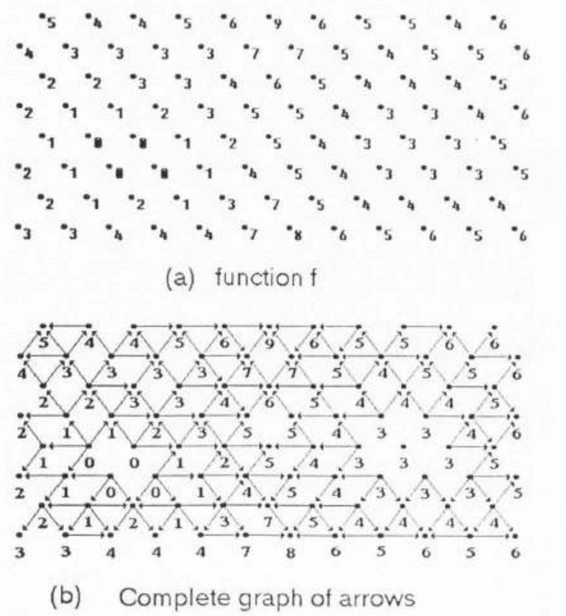


Figure 7. Function f (a) and its complete graph of arrows (b).

the end of the process, the watershed line DL(f) is equal to:

$$DL(f) = W^c_N(f) \quad \text{(with } \max(f) = N) \tag{15}$$

<u>Other Algorithms</u>

The watershed algorithms can be divided in two groups. The first group contains algorithms which simulate the flooding process. The second group is made of procedures aiming at the direct detection of the watershed points.

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 briefly present another algorithm belonging to the second group and based on the *arrows representation* of a function f [1].

From $f : Z^2 \rightarrow Z$, we define an oriented graph whose vertices are the points of Z^2 and with edges or arrows from x to any adjacent point y if and only if $f(x) < f(y)$ (figure 7).

The definition does not allow the arrowing of the plateaus of the topographic surface. This arrowing, named *completion* of the arrows graph, can be performed by means of geodesic dilations which propagate the descending borders of the plateaus inside them. Moreover, in order to suppress problems due to the fact that a watershed line is not always of zero thickness, a more complicated procedure called *over-completion* is used, which leads to a double arrowing for some points.

Then, starting from this complete graph (over-completed), we may select some configurations which, locally, correspond to divide lines. These configurations are represented on figure 8 for the 6-connectivity neighborhood of a point on a hexagonal grid (up to a rotation).

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

Figure 9. Watershed by arrowing: primary divide points (a), final result (b).

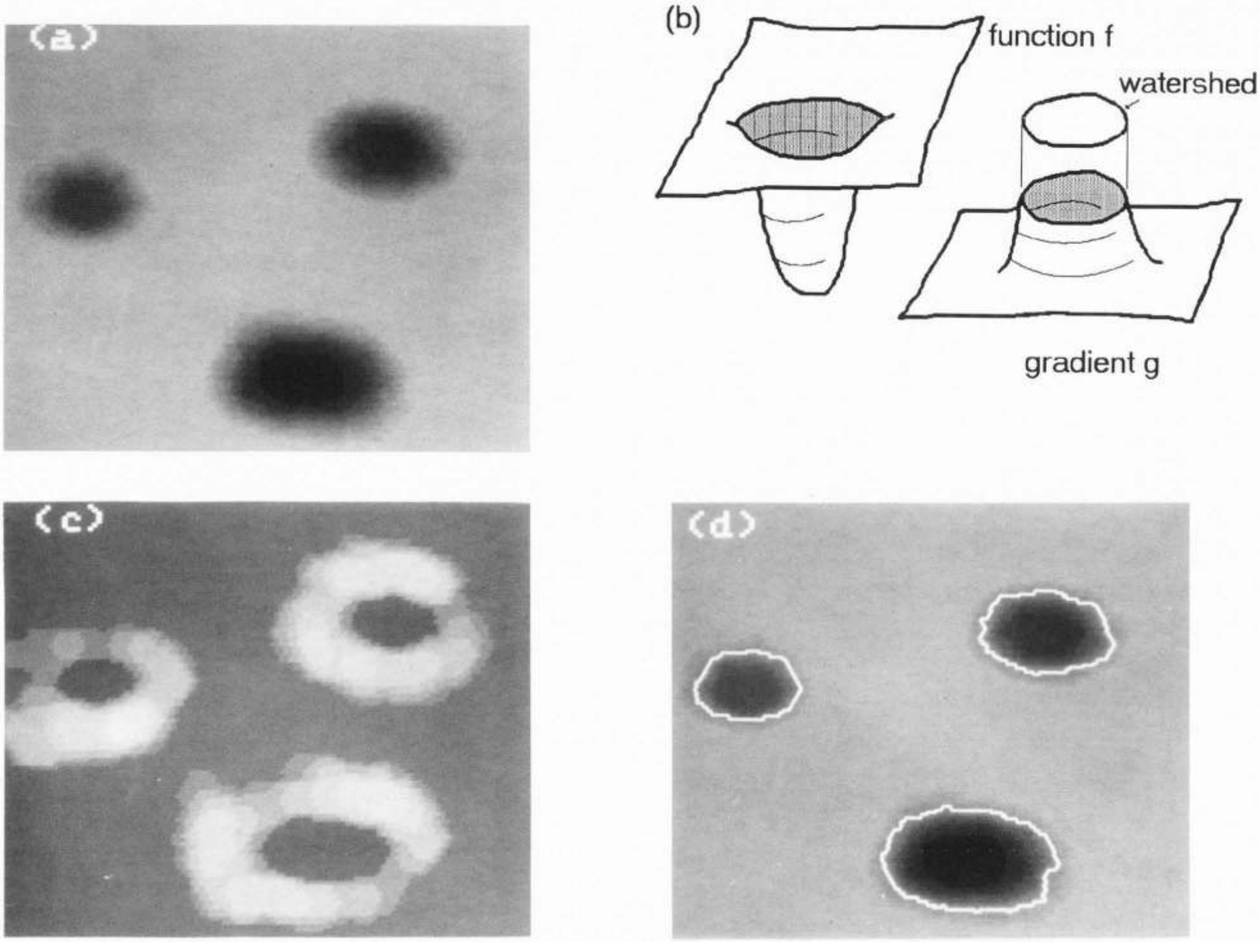


Figure 10. Simple blobs in a radioactive material (a), topographic surface of the initial function and of the gradient image (b), morphological gradient (c), watershed transform of the gradient image (d).

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 neighboring points and consequently the graph of arrows. Provided that the over-completion of this new graph has been made, some new divide points may then appear. The procedure is re-run until no new divide point is selected (figure 9).

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

Many watershed algorithms exist. They aim at reducing the computation time by only taking into account the points in the image that need to be modified at each step of the process. These algorithms are detailed in [6],[8].

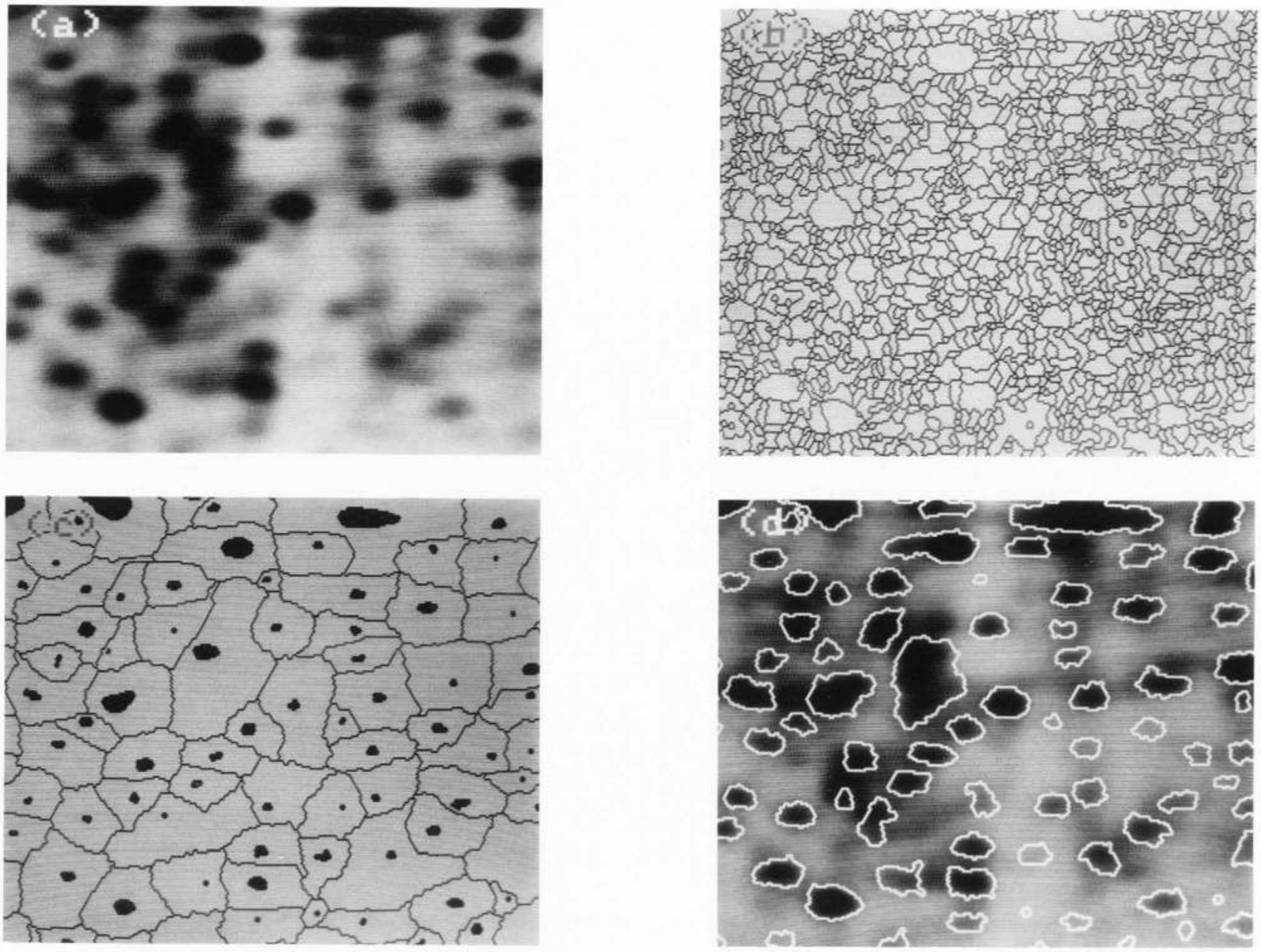


Figure 11. Electrophoresis gel (a), watershed of the gradient image (b),set of selected markers (c), final segmentation (d).

Application to Image Segmentation

Principle

The application of the watershed to image segmentation will be explained through a didactic example: the segmentation of single dots in an image (radon gas bubbles in a radioactive material).

The dots in figure 10a appear as domes with a round summit. Each dome has a unique summit. Our problem is to find the best contour.

A solution consisting of simply using a threshold is not sufficient because with a low threshold, the lowest domes are correctly detected, but the highest domes are much too large. A higher threshold, while detecting correctly the higher domes, misses the lower.

Since absolute values cannot be used, we may try instead the variation of the function, that is its gradient (figure 10c). The corresponding gradient image should present a volcano-type topography as depicted in figure 10b. The contours of the protein blobs correspond therefore to the watershed lines of the gradient image g(f) (figure 10d). In the new image, each dot of the original image becomes a regional minimum surrounded by a closed chain of mountains, like a basin. The varying altitude of the chain of mountains expresses the contrast variation along the contour of the original dot.

The over-segmentation problem

We can try to solve a similar problem, the contouring of proteins in an electrophoresis gel, by the same procedure (figure 11).

Unfortunately, the real watershed transform of the gradient, given in figure 11b, present many catchment basins. Each catchment basin corresponds to a minimum of the gradient. These minima are produced by small variations, mainly due to noise, in the grey values. This over-segmentation could be reduced by appropriate filtering. But a better result would be obtained if we mark the patterns to be segmented before performing the watershed transformation of the gradient. Suppose that we mark each blob of protein of the figure 11a. 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, we apply the watershed to the initial image. Then, we obtain a set of markers M (figure 11c). We consider again the topographic surface of the gradient

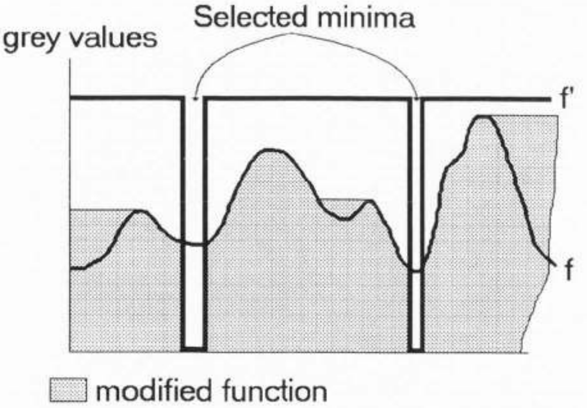


Figure 12. Principle of the homotopy modification of a function f by a set of selected minima.

image and the flooding process, but, instead of piercing the minima of this surface, we 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 in the marker set. Moreover, the watershed lines corresponding to the contours of the objects will occur on the crest lines of this topographic surface (figure 11d).

This algorithm can be written as follows. If $W_i(g)$ is the section at level i of the new catchment basins of g, we have:

$$W_{i+1}(g) = IZ_{Z_{i+1} \cup M}(W_i(g)) \qquad (16)$$

with:

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

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

Homotopy modification

The previous procedure can be implemented in 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. We have:

$$\forall i,\ Z_i(g') = R_{Z_i(g) \cup M}(M) \qquad (17)$$

This transformation is called *geodesic reconstruction* of a function. The gradient function g controls the reconstruction of a function defined from the markers M as illustrated in figure 12.

The second step simply consists in performing the watershed of the modified gradient g'.

The Segmentation Paradigm

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' (figure 13).

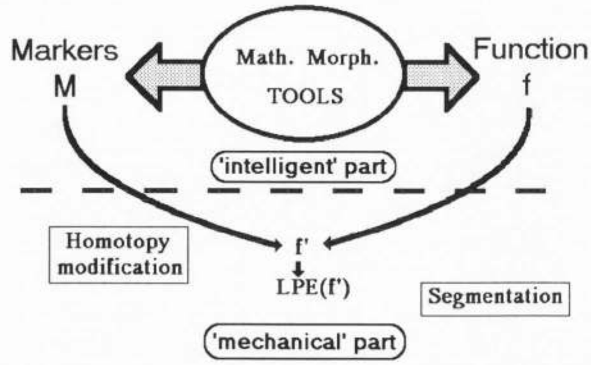


Figure 13. Synopsis of the morphological segmentation methodology.

The segmentation process is therefore divided in 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 namely: watershed and image modification.

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

Segmentation of overlapping grains

The figure 14a represents a transmission electron microscopy (TEM) image of grains of silver nitrate scattered on a photographic plate. Some of them are overlapping and they need to be segmented in order to measure without bias their size and shape.

To apply the methodology described above, the background, the grains and the overlapping regions must be pointed out. To do so, we first threshold the initial image (an automatic thresholding can be performed without difficulty) (figure 14b). Then, the maxima of the distance function d(X) of the binary image X provide the markers of the grains (figure 14c). The markers of the overlapping regions are obtained in a more refined way. The watershed transformation of the inverted distance function -d(X) produces divide lines which cut the overlapping grains (figure 14d). These divide lines pass through the overlapping regions and consequently are used to mark them. These markers correspond to the centers of the divide lines (figure 14e). The marker of the background is simply the set X slightly eroded (figure 14f).

The function controlling the segmentation is the gradient function (figure 14g). The homotopy modification and the watershed construction are performed. The figure 14h shows the final result, after the elimination of the artifacts.

Stereoscopic analysis of a fracture in steel

The second example is a problem of segmentation of cleavage facets in a SEM micrograph of a steel fracture (figure 15). The function used for the watershed along with the markers set 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).

Two functions are defined: the first one, f_1, is the supremum of the gradient function of the initial image f and of a morphological transformation called *"Top-Hat" transformation* [5]. The Top-Hat transform TH(f) defined as the

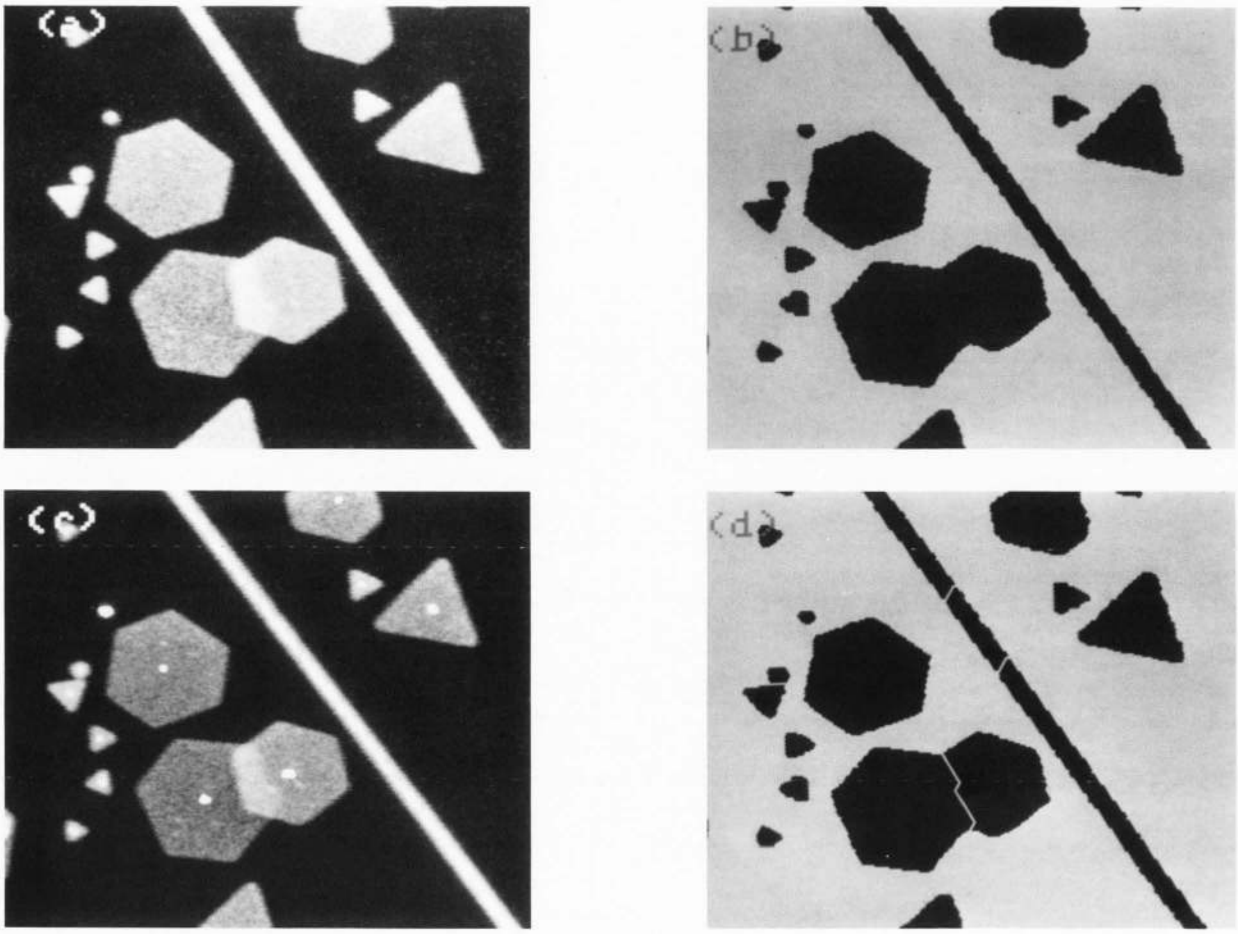


Figure 14. TEM image of silver grains (a), thresholded image of grains (b), markers of the grains (c), first segmentation of the grains (d).

difference between the function and its morphological opening is a contrast detector suitable for enhancing in the image the blazing zones (figure 16a):

$$f = \mathrm{Sup}\,(g(f),TH(f)) \qquad (18)$$

The second function f_2 is the distance function to the blazing zones and to the contours. It can be shown [1] that this function may be built by dilating the previous function f_1 by a cone (figure 16b).

The markers of the facets are the minima of f_2 (figure 16c). We can see that more than one marker may appear in regions which obviously correspond to simple facets. This multiple marking leads to an over-segmentation of the facets.

In order to eliminate this over-segmentation, the watershed transformations of the two functions f_1 and f_2 are performed (figure 16d) and only the divide lines which are superimposed in the two watershed transforms are kept (figure 16e).

The methodology of the segmentation based on the primary definition of the markers of the objects to be extracted is particularly helpful here. Indeed, when the first picture of the stereoscopic pair has been segmented and the corresponding facets selected, the markers used in this first step can be used again to segment the homologous facets in the second picture of the stereo pair. The procedure is the following: the markers attached to a facet in the first image are "thrown" onto the second image f_2' corresponding for the second picture to the image f_2 . These markers fall along the steepest slope of f_2' and each one reaches a unique minimum of f'_2. These minima are the markers of the homologous facet in the second picture (figure 17). Doing so, we establish a one-to-one correspondence between the markers of the two pictures of the stereo pair and therefore, between the segmented facets (figure 18).

As soon as the same facet (or part of a facet) has been segmented in the two pictures of the stereo pair, the computation of its size and orientation in space is relatively easy. By following the corresponding points in the two contours, it is possible to calculate the shift between them and hence their height. Assuming that a facet is almost a plane, its interpolation is performed. Finding the cleavage angle between two adjacent facets (which is in fact the required parameter) is immediate.

This approach of the stereovision consisting in first segmenting the objects instead of trying to find immediately

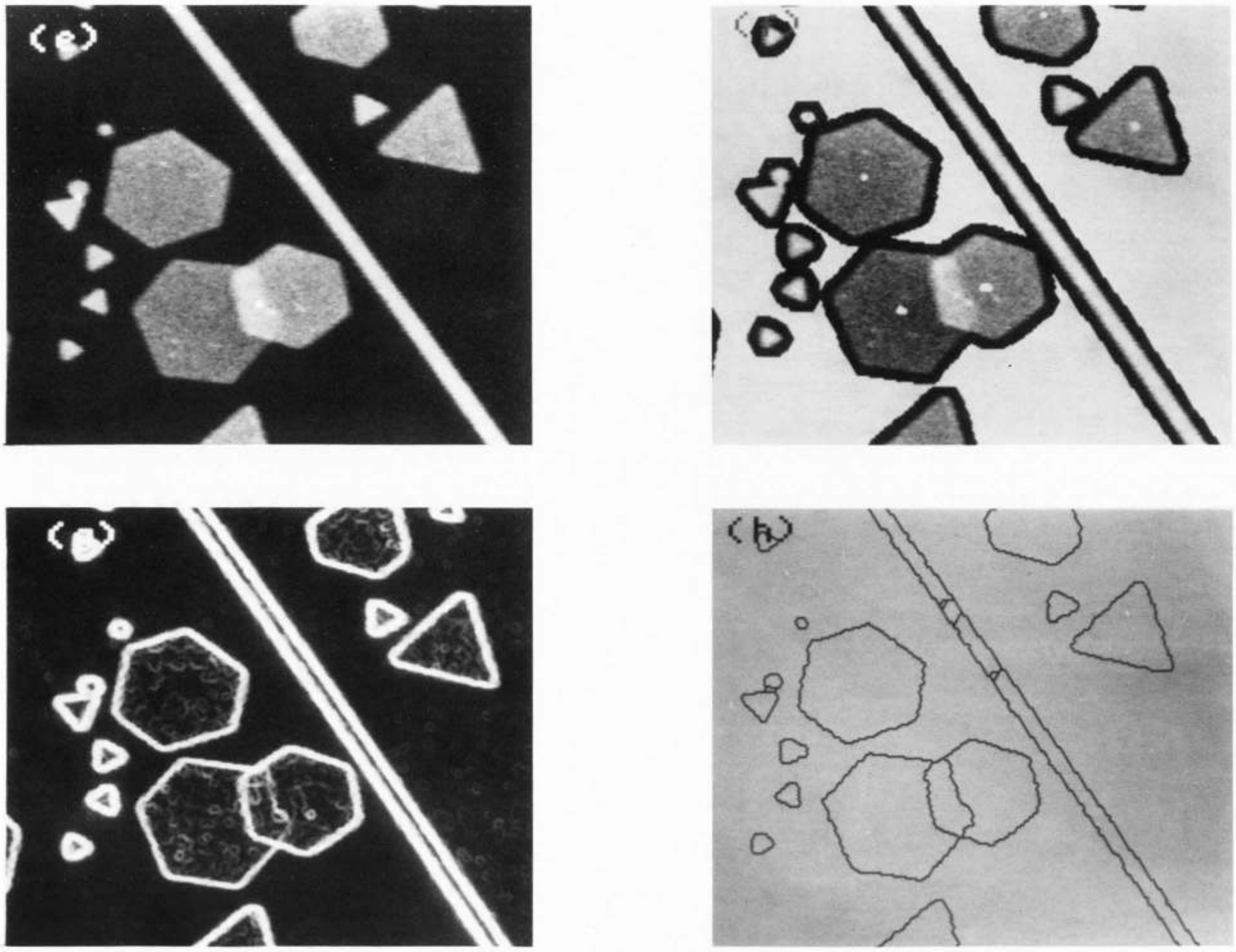


Figure 14 (continued). Final markers of the overlapping regions (e), set M of markers (f), gradient image (g), final segmentation (h).

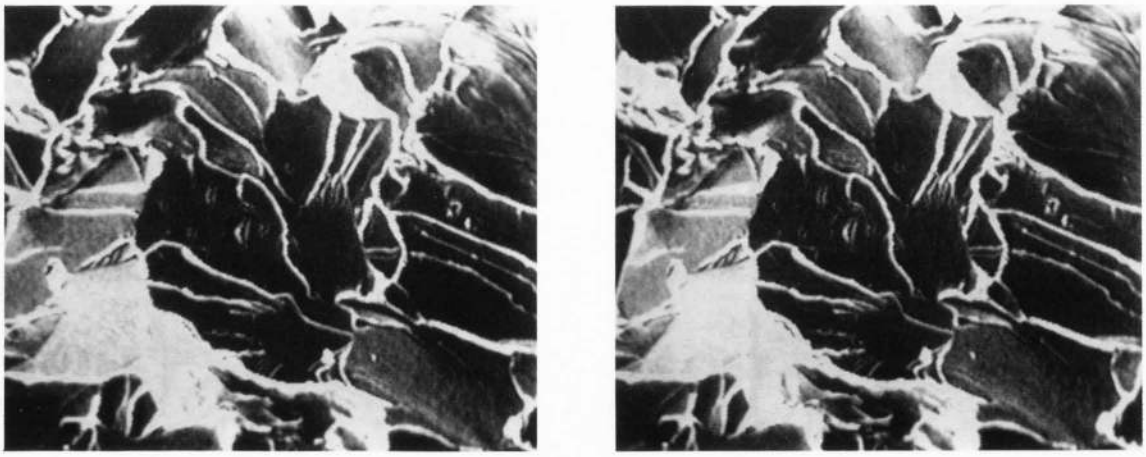


Figure 15. Stereo pair of a cleavage fracture in steel.

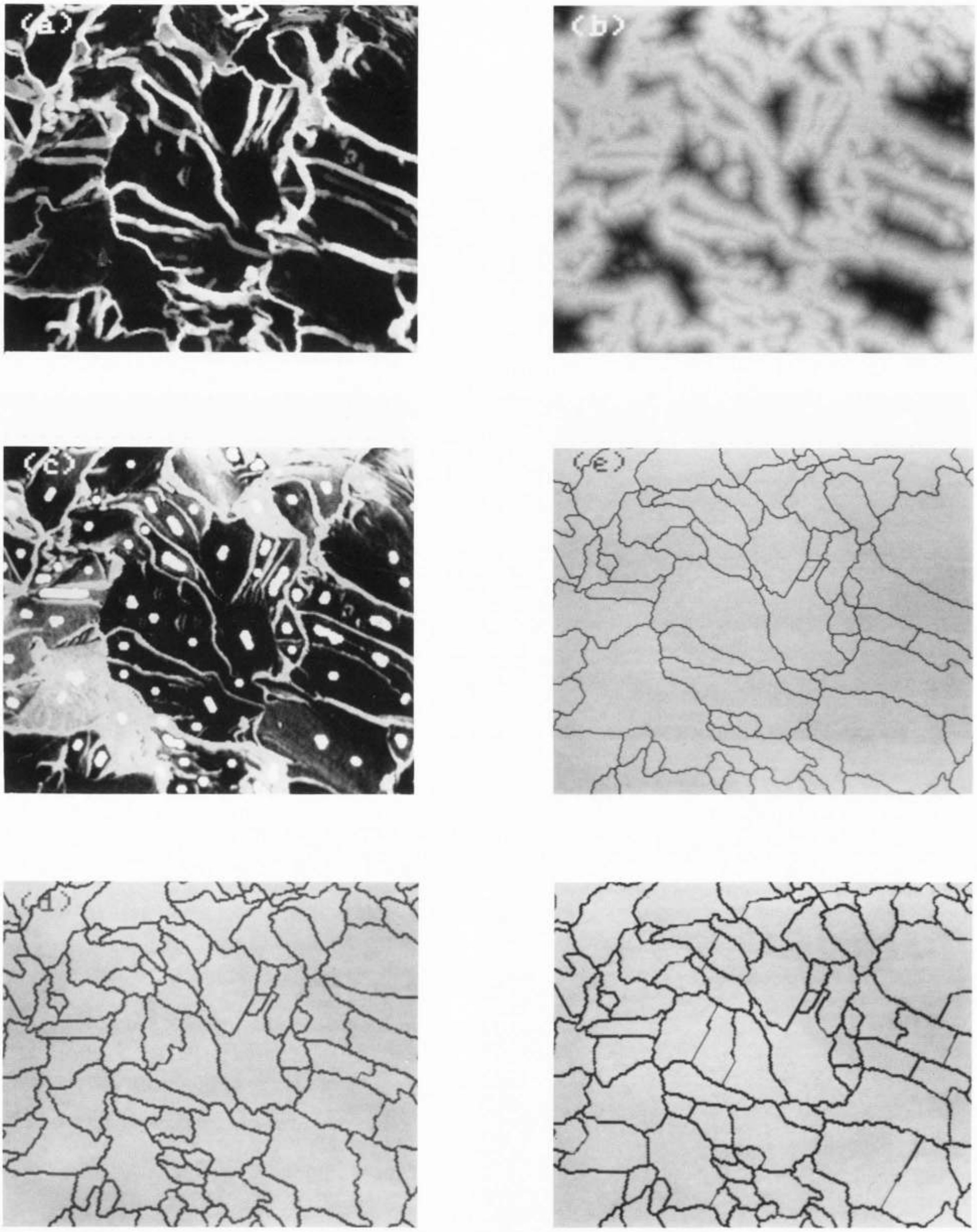


Figure 16. First function used for segmentation (a), second function (b), markers of the facets (c). Watershed lines of the two functions f and f (d), final contours of facets (e).

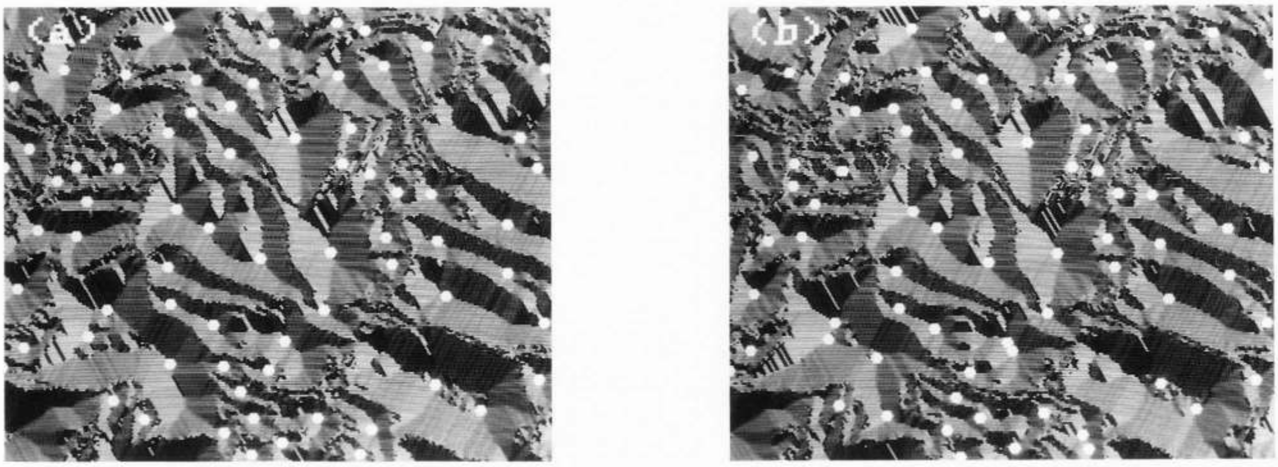


Figure 17. Markers of the first image (a), corresponding markers in the second one (b).

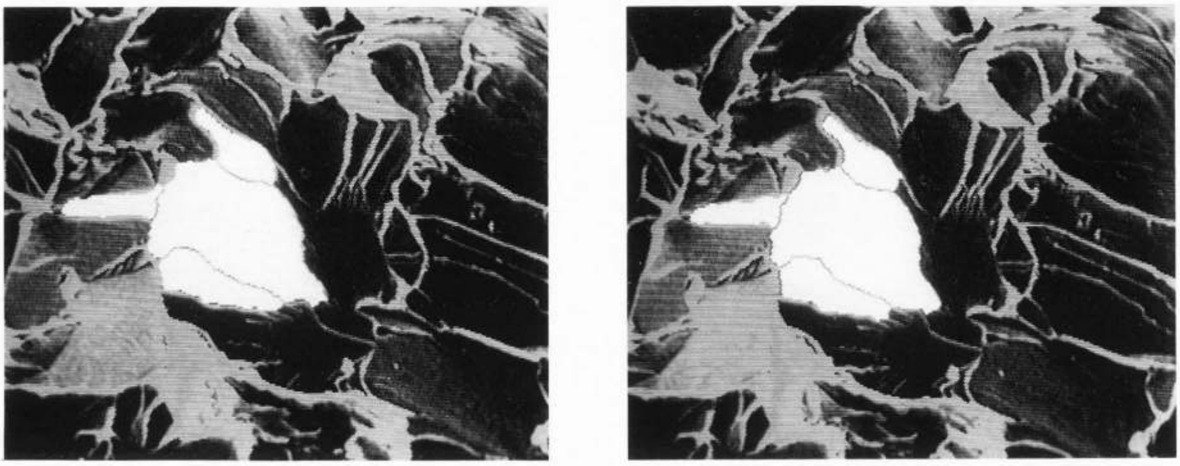


Figure 18. Homologous facets in the stereo pair.

the homologous pixels in the two images is very powerful: the watershed transformation coupled with the markers selection allows us to directly find the corresponding objects in the stereo pair. Moreover, this topological approach allows the very accurate control of this correspondence (two adjacent objects in the scene are in most cases adjacent in both images of the stereo pair).

Hierarchical Segmentation

Introduction

Unfortunately, in some cases, the markers selection and extraction are not so easy. Some pictures may be very noisy and image processing becomes more and more complex. In other cases, the objects to be detected may be so complex and so varied in shape, grey level and size that it is very hard to find reliable algorithms enabling their extraction. For that reason, we need to go a step further in the segmentation.

We know that the initial watershed transformation of the gradient image provides very unsatisfactory results: many apparently homogeneous regions are fragmented in small pieces. Fortunately, the watershed transform itself, applied on another level, will help us to merge the fragmented regions. Indeed, if we look at the boundaries produced by the segmentation, they do not have the same weight. Those which are inside the almost homogeneous regions are weaker. In order to compare these boundaries, we need to introduce neighborhood relations between them through the definition of a new graph. This graph is built from a simplified version of the original image called partition or *mosaic image*.

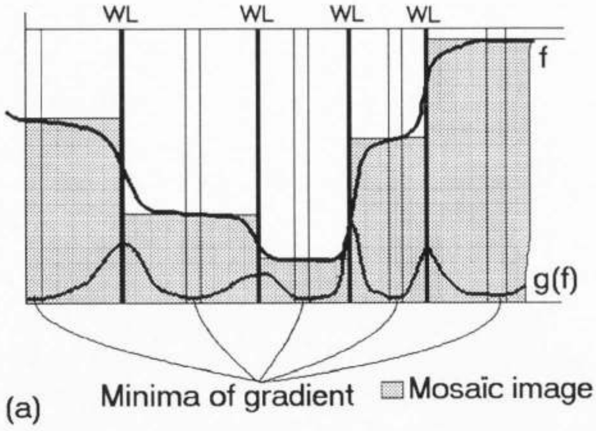
The mosaic image

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

The figure 19a illustrates this operation. The initial image is an X-ray photograph of metallic particles in the burst produced by a shaped charge weapon. The result is a simplified image (figure 19b), made of a mosaic of pieces (the catchment basins) of constant grey levels, where no



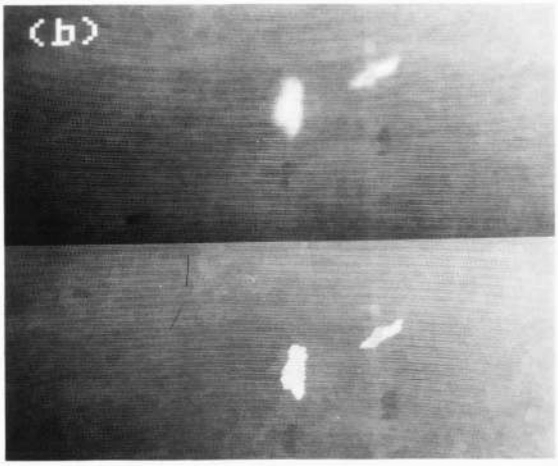


Figure 19. Computation of the mosaic image (a), initial (upper) and mosaic (lower) images (b).

information regarding the contours has been lost. This simplified image, also called mosaic image, may then be used to define a valued graph, to which the morphological transforms, and in particular the watershed, can be extended.

Hierarchical segmentation

Let us build a new valued graph from the mosaic image. First, two boundaries of the mosaic are considered neighbors if they surround the same catchment basin. Second, the boundaries between two tiles of the mosaic image are valued with the grey tone difference between these tiles.

All the morphological transformations can be extended to the resulting graph illustrated in figure 20, where the summits correspond to the simple arcs of the primitive watershed transform and the vertices connect the boundaries surrounding the same primitive catchment basin. In particular, the notion of minimum as it is defined above using paths on the graph of a function, can be applied to this valued graph. In our case, the weakest boundaries of the mosaic image correspond to regional minima of the new graph (figure 21a). We may flood the relief of the graph starting from these minima. All the boundaries inside the catchment basins are suppressed. Only the boundaries corresponding to the divide lines of the graph remain. Doing so, we have suppressed the boundaries of the primitive watershed which are surrounded by the more contrasted ones. The result of this hierarchical segmentation is given in figure 21b. From that picture, the extraction of the particles is straightforward. They correspond to the new catchment basins that contain the maxima of the initial image (figure 21c).

Other examples

This hierarchical segmentation can be used efficiently for extracting features from complex scenes. For instance, this technique has been applied for delineating the road in figure 22.

The result of the watershed transformation yields to a hierarchical segmentation of the image, as illustrated in the previous example. The selection of some 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 [2].

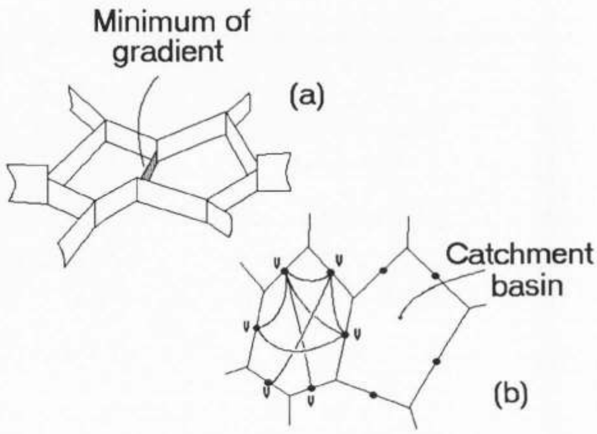


Figure 20. Gradient of the mosaic image (a), the corresponding graph used in the hierarchical approach (b).

Starting from a highly fragmented image, we have obtained a new mosaic after simplification. It is obviously possible to iterate this simplification process. By this means we get a hierarchy of simplification stages, the last always being a uniform image.

Conclusions

The segmentation of images by means of the watershed transform and the use of markers has many advantages:

- The watershed transform provides closed contours by construction.
- When computing the watershed, there is a good match between the contours which undoubtedly appear in the image and the divide lines of the gradient watershed, even when it is severely over-segmented.
- It is a general method which can be applied in many situations. The examples given in this paper are in fact a small selection of the domains in image analysis where this technique has been used efficiently.

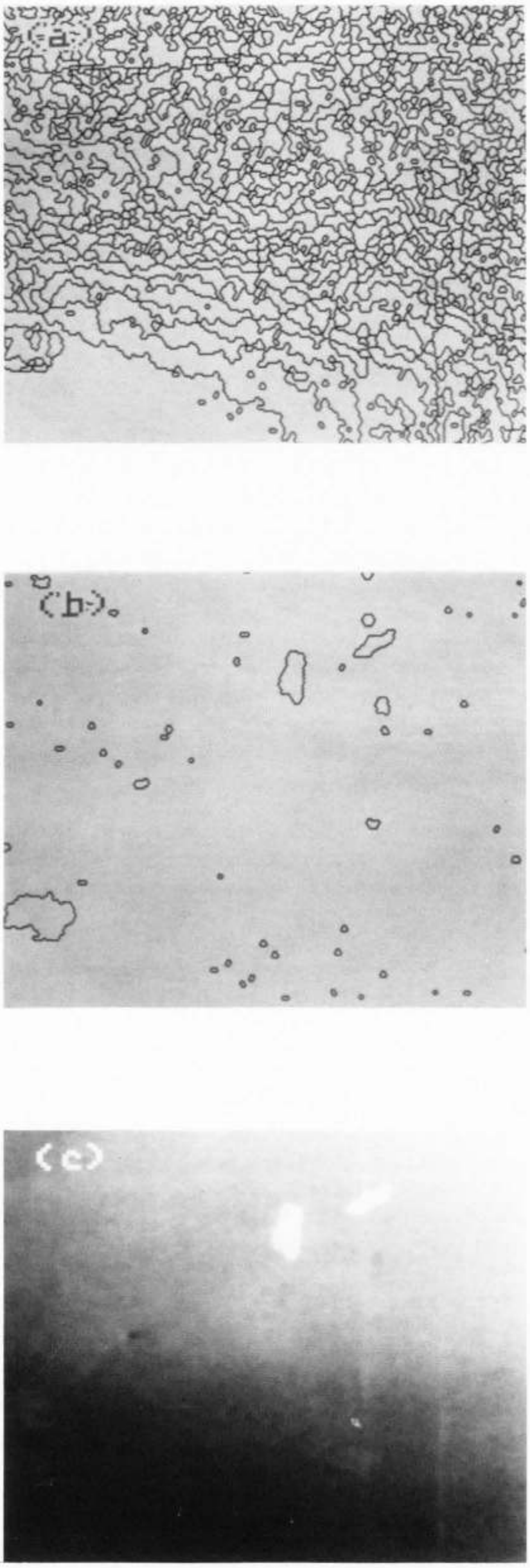


Figure 21. Initial watershed (a), hierarchical segmentation (b), final result (c).

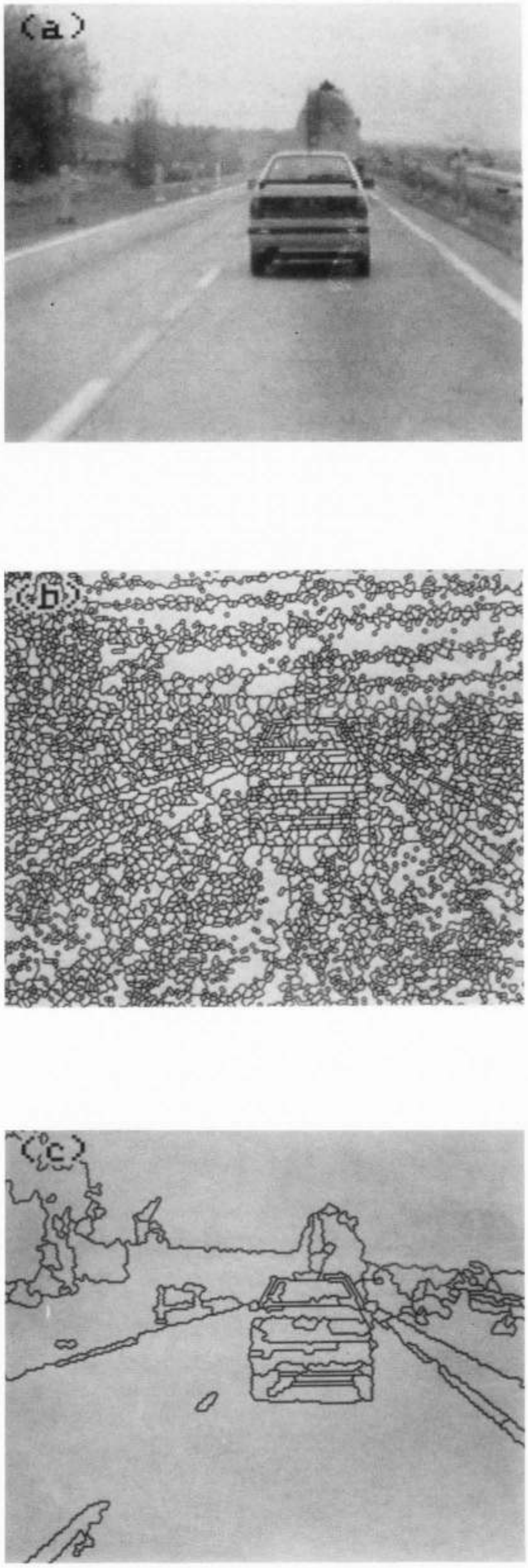


Figure 22. Road scene (a), first level of segmentation (b), second level of the hierarchy (c).

- The great advantage of this methodology is that it splits into two separate steps the segmentation process. First, we must detect what we want to extract: it is the markers selection. Then, we define the criteria required to segment the image.

This last assertion means that image segmentation cannot be performed accurately and adequately if we do not construct 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. Beucher S (1990) *Segmentation d'Images et Morphologie Mathématique (Image Segmentation and Mathematical Morphology)*, Doctorate Thesis, Paris School of Mines, France, pp. 111-273, June 1990.

2. Beucher S, Bilodeau M and Yu X (1990) Road Segmentation by watershed algorithms. Proc. of PROMETHEUS workshop, Sophia-Antipolis, France, pp. 8-31, April 1990.

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

4. Lantuéjoul C, Beucher S (1981) On the use of the geodesic metric in image analysis. J Microsc 121, 39-49, Jan. 1981.

5. Meyer F (1979) *Cytologie Quantitative et Morphologie Mathématique (Quantitative Cytology and Mathematical Morphology)*, Doctorate thesis, Paris School of Mines, Paris, France, pp. 101-106, May 1979.

6. Meyer F (1989) Skeletons and perceptual graphs. Signal Process 16, 4, 335-363, April 1989.

7. Meyer F, Beucher S (1990) Morphological Segmentation. JVCIR 1, 1, 21-46, Sept. 1990.

8. Soille P, Vincent L (1990) Determining watersheds in digital pictures via flooding simulations. Proc. SPIE Visual Communications and Image Processing '90, 1360, 240-250, Oct. 1990.

Discussion with Reviewers

Reviewer: The tools used in the paper are sufficiently varied and flexible that a variety of segmentation tasks can be performed by applying them in various combinations. But little is said about the assumptions under which the results could be expected to be satisfactory, and nothing (quantitative) is said about the assumptions that must be satisfied by the (ideal) image and the noise. Tools are nice, but if there are no probabilistic models or model-based criteria for tools selection, the approach lacks a firm theoretical foundation.

Author: The answer to this comment is more or less contained in the diagram of figure 13. The tools described in this paper are used in the second part of the segmentation process (the "straightforward" one). To segment correctly objects in a picture, you must first understand what you are looking at and build the objects you want to extract. Understanding an image is necessarily the first step of the segmentation. This step is formalized through the definition of the markers and the selection (and even the construction) of the function which will be used for the watershed transformation. This "intelligent" level mimics what we do when we look at a picture: we point out the objects of interest (markers selection) and we explain which criterion is used for distinguishing the relevant features (criterion function construction). This criterion function may be the gradient, but also other functions not related to contrast variations. There is no general scheme to achieve this first step because there is no general representation or model for describing all the objects you can see in the real world. You must use, in this part of the process, the most efficient tools adapted to the problem you are dealing with, and, more especially, probabilistic models or model-based criteria. Note that the sole purpose of image segmentation is to quantify (size, shape, etc.) the features under study. We are interested mostly in an efficient technique (even if it can be considered as "ad hoc") able to process many samples in a reasonable amount of time rather than in a general tool that could be used whatever the kind of picture being analyzed.

Moreover, the watershed transformation coupled with the marker selection and the homotopy modification is dramatically insensitive to the noise in the image.

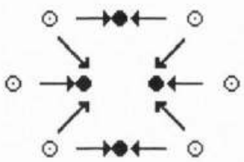
Finally, I do not believe that a tool needs a "firm theoretical foundation" to be used and to be performing.

N.K.Tovey: Presumably each of the marker blobs in Fig. 11c should be contained within a single area as defined by the lines in Fig. 11b (otherwise some over-segmentation lines will still remain). In the figures as presented, it appears that some blobs may cross lines. Is this an optical illusion? Please clarify this point.

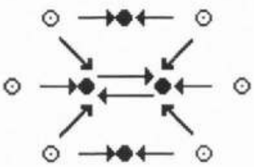
Author: It is not an optical illusion. The markers may cross the watershed lines of the gradient image. In fact, there is no relationship between the markers depicted in figure 11c and the original catchment basins of the gradient. The figure 11b is just given to illustrate the over-segmentation of the gradient watershed which occurs when no marker is used to point out the objects to be segmented. The entire process follows the steps given in figure 13: first, we select the markers of the objects to be segmented (figure 11c), then we modify the homotopy of the gradient image. Finally, the watershed transformation of the modified gradient image is performed to produce the contours of the blobs (figure 11d). Note that the contours obtained in this last figure are a subset of the watershed lines of the original gradient (figure 11b).

J.C.Russ: The completion of the arrows graph is described nicely but the "more complicated procedure called over-completion" which is actually used, is not shown or discussed.

Author: The completion of the arrows graph is the step of the arrowing process which allows to take into account the plateaus (flat zones) of the topographic surface. When propagating the flood, the completion is the part of the process which defines how the flow invades the plateau: it starts from the ridges and fills progressively the plateau (exactly as a geodesic dilation would do). The over-completion is a trick to solve a very annoying problem encountered with digitized pictures: the parity problem. This problem clearly appears in the following example:



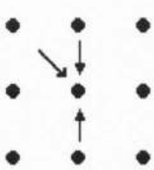
Despite the fact that the four black points obviously belong to the watershed line, the two central ones will never be selected: the arrows graph is complete and the neighborhoods of the latter points do not belong to the configurations given in figure 8. The over-completion solves this problem by adding double arrows between the two central points:



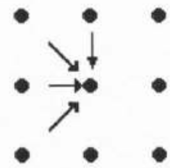
Therefore, their neighborhood can be selected. It can be shown [1] that completion and over-completion can be performed by using the same rules.

J.C.Russ: The methods described are specially for the case of hexagonal pixels, whereas most image analysis systems in fact use square pixels. Modification of the neighbor counting rules for square pixels would be very helpful.

Author: The configurations which may correspond to divide lines have been described for the hexagonal grid for the sake of simplicity. In this case, indeed, the neighborhood of any point is unique: it is made of six neighbors. In the case of the square grid, two connectivity relationships can be used: the 4-connectivity or the 8-connectivity. Moreover, these connectivities must be different for the "objects" and for the background. This means that, for the watershed transform, if you decide to deal with 8-connected catchment basins, the watershed lines will be necessarily 4-connected. As soon as the connectivity relationships are established, the rule for finding the right configurations is easy: when there are at least two connected components of arrows, this configuration corresponds to a divide line. For instance, with 8-connected catchment basins, in this configuration:



the central point belongs to a divide line. Conversely, this configuration:



will not be retained.

N.K.Tovey: You refer to "a given neighborhood" (in the description of the morphological gradient). What size is this neighborhood region?

Author: This neighborhood corresponds most of the time to the elementary structuring element. But larger structuring elements can be used to produce what is called "thick gradients". These thick gradients are mainly used for defining more sophisticated gradients called "regularized morphological gradients" which are very efficient on noisy images [1].

N.K.Tovey: Presumably it is only possible to get true separation of the two overlapping particles in the center of Fig. 14a if there is sufficient information from the lower particle to define the "hidden" edge with the gradient operator. If the upper particle is thick the segmentation by this method would not seem possible. Is this correct?

Author: Yes, it is true. Here again, to segment the different "objects" or regions in the image, a criterion must be defined. In this example, the variation in grey values is the used criterion. We know, because we work with a TEM image, that there is no hidden part when grains overlap (the grains are transparent). If there were some hidden regions in the image, you should introduce further knowledge about the shape of the hidden parts. This information obviously could not be extracted from the images.

N.K.Tovey: When the "markers" are "thrown" onto the second image (in the fractures example), the markers are allowed to roll to the relevant minimum in the second image. Do difficulties arise if a facet in the first image is steep and the corresponding facet is either hidden or vertical in the second image?

Author: Yes, this technique assumes that there is a one-to-one correspondence between the facets of the first image and the homologous ones in the second picture. As you mentioned, it may happen that some steep facets in the first picture disappear in the second one. For this reason, the real procedure is more refined. First, instead of "throwing" at the same time all the markers on the second picture, we start with the markers which are close to the tilt axis (the vertical axis in the middle of the image). Secondly, each marker of the first image is translated according to its distance from the tilt axis before it is thrown. This step is needful to correct the fact that, when you tilt a sample, the markers are moving. Finally, a similarity test is performed between the facets supposed to be homologous in the stereo pair. In particular, these facets should have almost equal vertical dimensions.

N.K.Tovey: Reference is made to computation of spatial orientation of facets. Such computation is only correct if the precise geometry of the two images is known (e.g. 3 rotations + 3 translations). These elements of interior and exterior orientation must be known. Were these elements determined or are the computations using approximate photogrammetric formulae?

Author: Some geometric elements are already known, in particular the magnification and the tilt angle. To determine the orientation of each facet, we calculate the shifts between the corresponding points of its boundary on the two pictures of the stereo pair and deduce their height by a photogrammetric formula. Each facet is supposed to be more or less flat, so we can estimate the orientation and steep of the best fitted plane passing through the boundary points. Moreover, a dispersion calculus allows to check if the assumption of a plane facet is correct or not.