Geodesic Saliency of Watershed Contours and Hierarchical Segmentation

Laurent Najman, Michel Schmitt

Geodesic Saliency of Watershed Contours and Hierarchical Segmentation

Laurent Najman and Michel Schmitt

Abstract—The watershed is one of the latest segmentation tools developed in mathematical morphology. In order to prevent its oversegmentation, the notion of dynamics of a minimum, based on geodesic reconstruction, has been proposed. In this paper, we extend the notion of dynamics to the contour arcs. This notion acts as a measure of the saliency of the contour. Contrary to the dynamics of minima, our concept reflects the extension and shape of the corresponding object in the image. This representation is also much more natural, because it is expressed in terms of partitions of the plane, i.e., segmentations. A hierarchical segmentation process is then derived, which gives a compact description of the image, containing all the segmentations one can obtain by the notion of dynamics, by means of a simple thresholding. Finally, efficient algorithms for computing the geodesic reconstruction as well as the dynamics of contours are presented.

Index Terms—Morphological segmentation, watershed, dynamics, hierarchical segmentation, geodesic reconstruction.

1 Introduction

Segmentation and contour extraction are key points of image analysis. There are numerous algorithms for doing these operations, which have the drawback of producing an oversegmentation. Several techniques have been developed to diminish this oversegmentation, the most common one being **hysteresis thresholding** by Canny [5]. Combined with the noise reduction induced by the Gaussian convolution, it has largely contributed to the success of Canny's extractor.

Mathematical morphology uses the **watershed** algorithm, introduced for the purpose of segmentation by Lantuéjoul and Beucher [3], and mathematically defined in [14], [16]. See [12] for a definition from a more algorithmical point of view. As the other techniques, the watershed produces an oversegmentation and until now, a procedure similar to "hysteresis thresholding" has not existed. The aim of this paper is to introduce such a procedure.

Watershed is often used in conjunction with **geodesic reconstruction**, a powerful tool developed by mathematical morphology, which simplifies gradient images and prevents oversegmentation. In this paper, we present a new algorithm which aims at computing in one step all the segmentations by watersheds that one can obtain by the use of geodesic reconstruction, or, equivalently, by the concept of **dynamics** [9]. The main advantage of our algorithm is that it directly gives a **hierarchical segmentation** in which all the contour arcs are evaluated by a measure of saliency (and not the catchment basins, as originally proposed by Grimaud), allowing one to choose the desired level of dynamics after the segmentation process. This measure of the saliency of the contour arcs will also be called "dynamics," because it is based on the same morphological tool, namely the geodesic reconstruction. The result can then be used in a similar way to the hysteresis thresholding, but in the context of region based segmentation.

This paper presents the state of the art around ideas of morphological segmentation, while mainly focusing on the geodesic reconstruction, and placing these ideas in a new unifying perspective which leads to the novel concept of dynamics of contours. It is organized as follows: First, we review the basic definition of the watershed. We then present the usual ways to reduce oversegmentation, which all rely on the geodesic reconstruction. Then we introduce the principle of hierarchical segmentation. Under this framework, we present a new concept of dynamics of contours, which allows a valuation of watershed contours relying on the gradient information and on the geodesic reconstruction. Then we discuss the interest of our concept, both from a mathematical and a practical point of view, illustrated by an application to shape recognition. Finally, we propose an algorithm to compute this hierarchical segmentation efficiently, together with a novel algorithm to compute the geodesic reconstruction.

2 The Watershed: A Tool for Segmentation

This section presents the standard definition of the watershed and can be skipped by the reader familiar with mathematical morphology (see for instance [21]).

In mathematical morphology, it is usual to consider that an image is a topographical surface. This is done by considering the gray level (the image intensity) as an altitude. Places of sharp changes in the intensity thus make a good set in which one can search for contour lines. It is then rather straightforward to estimate the variation from the gradient of the image. For the purpose of segmentation, we

- *L. Najman is with ArSciMed, 207 rue de Bercy, 75012 Paris, France. E-mail: lnajman@arscimed.com.*
- *M. Schmitt is with the Centre de Geostatistique, Ecole des Mines de Paris, 35, rue Sant-Honore 77305 Fontainebleau-Cedex, France. E-mail: mschmitt@cg.ensmp.fr.*

Manuscript received Nov. 10, 1994; revised Aug. 10, 1996. Recommended for acceptance by G. Medioni.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96084.

then look for the **crest lines of the gradient image**. A way to characterize these lines is to apply the watershed algorithm to the modulus of the gradient image.

The idea of the watershed [6], [3] is to attribute an **influence zone** to each of the **regional minima** of an image (connected plateau from which it is impossible to reach a point of lower gray level by an always descending path). We then define the watershed as the boundaries of these influence zones.

Numerous techniques have been proposed to compute the watershed. The major ones are reviewed in [23], [25]. The classical idea for building the watershed is illustrated in one dimension (Fig. 1). Using a geographical analogy, we begin by piercing the regional minima of the surface, then slowly immerse the image into a lake. The water progressively floods the basins corresponding to the various minima (Fig. 1a). To prevent the merging of two different waters originating from two different minima, we erect a dam between both lines (Fig. 1b). Once the surface is totally immersed, the set of the dams thus built is the watershed of the image. In one dimension, the location of the watershed is straightforward: it corresponds to the regional maxima of the function. In two dimensions (which is the case for gray-scale images), this characterization is not so easy. The place where two basins meet for the first time is a saddle point in the image. One can say in an informal way that the watershed is the set of **crest lines** of the image, emanating from the saddle points.

We present here the classical algorithm for computing the watershed, in the case of a function defined in $\mathbb{R}^2$ or on a digital grid, with discrete range (step functions). The most powerful implementation described in the literature ([25], [24], [22], [4]) uses FIFO breadth-first scanning techniques for the actual flooding.

Following the ideas mentioned above, the algorithm consists in flooding the various basins, and in keeping as the watershed the set of contact points between two different basins. In the case where this contact is on a plateau, we keep the (geodesic) middle line of this plateau. The watershed thus defined is of thickness one on the grid.

DEFINITION 2.1 *Let A be a set, a and b two points of A. We call* ***geodesic distance*** *$d_A(a, b)$ in A the lower bound of the length of the paths γ in A linking a and b.*

Let B be a set included in A. The geodesic distance $d_A(b, B)$ from a point b to the set B is defined as usual by $d_A(b, B) := \min_{c \in B} d_A(b, c)$.

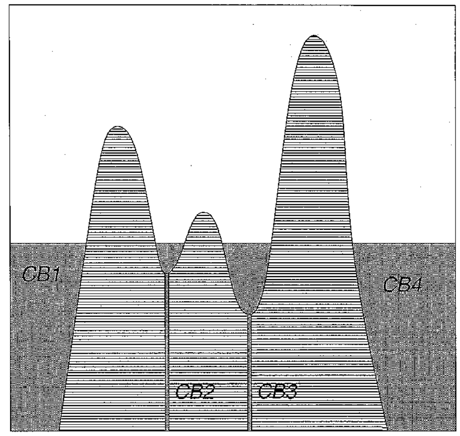
In the digital case, the distance d_A is deduced from the one on the grid [20].

Let $B = \cup B_i \subset A$, where B_i are the connected components of B.

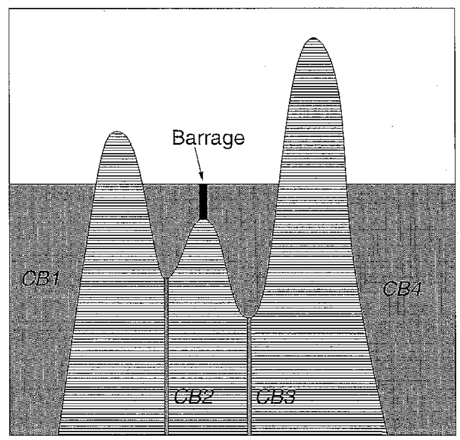
DEFINITION 2.2. *The* ***geodesic influence zone*** *$iz_A(B_i)$ of a connected component B_i of B in A is the set of the points of A for which the geodesic distance to B_i is smaller than the geodesic distance to other connected components of B.*

$$iz_A(B_i) = \{a \in A, \forall j \in [1, k] \setminus \{i\}, d_A(a, B_i) < d_A(a, B_j)\}. \quad (1)$$

The points of A which do not belong to any influence zone make up the skeleton by influence zone of B in A, denoted by $SKIZ_A(B)$:



(a) At time t, the dam is not yet constructed.



(b) At time $t + h$, we construct a dame to separate water from CB_2 and from CB_3

Fig. 1. Building of the watershed: one-dimensional approach.

$$SKIZ_A(B) = A \setminus IZ_A(B) \quad (2)$$

where $IZ_A(B) = \cup_{i \in [1,k]} iz_A(B_i)$.

The watershed algorithm on digital images by recurrence on the gray levels is [6]:

DEFINITION 2.3. *Let $f : \mathbb{R}^n \rightarrow \mathbb{N}$ be a bounded step function. We note*

1) $h_{min} = \min f$ *and* $h_{max} = \max f$,
2) $[f]^h$ *the upper threshold of f at level h* : $[f]^h := \{a \in \mathbb{R}^n \mid f(a) \leq h\}$,
3) $Reg\_Min_h(f)$ *the set of the regional minima of f at height h.*

The set of ***catchment basins*** *of f is the set $X_{h_{max}}$ obtained after the following recurrence:*

i) $X_{h_{min}} = f^{h_{min}}$

ii) $X_{h+1} = \mathrm{Re}\,g\_Min_{h+1}(f) \cup IZ_{[f]^{h+1}}(X_h), \ \forall h \in [h_{min}, h_{max} - 1]$.

The ***watershed*** *of f is the complementary of X_{max}* (Fig. 2).

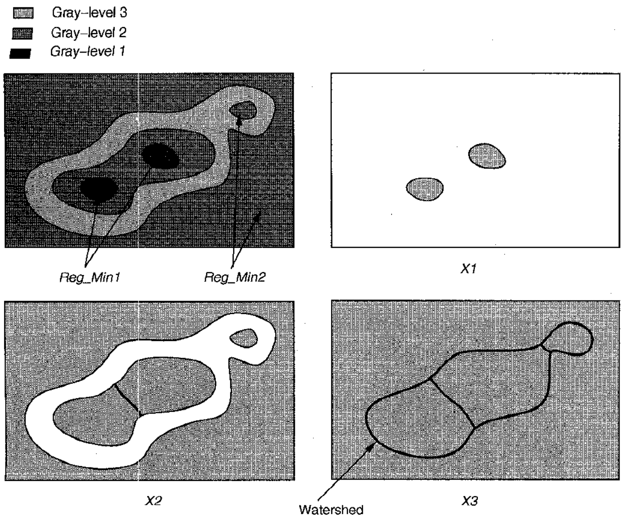


Fig. 2. Illustration of the recurrence immersion process.

3 Suppress the Oversegmentation

The watershed produces an oversegmentation of the images, but always contains contours which appear to be correct. The main problem is to make a choice between all those "right" contours. As in the case of Canny's extractor, the saliency of a contour can be evaluated by the value of the modulus of the gradient. But the step of hysteresis thresholding is not adapted to the watershed for three reasons:

1) Watershed produces a segmentation: Contours are obtained as complementary to the set of regions, and are consequently closed. Hysteresis thresholding is to be applied on edges and usually produces nonclosed contours. In other words, we start with a segmentation and get edges which do not necessarily build a segmentation.
2) Hysteresis thresholding on watershed segmentation produces barbs, which are small edges from adjacent regions (see Fig. 3). Only complicated algorithms could reliably eliminate these barbs.
3) The most important reason is probably that hysteresis thresholding is not a morphological process: It relies on the (local) neighborhood of the pixel, and not on the structure of the image. Hysteresis thresholding is well adapted to a local edge detector like Canny's one, but not to watershed segmentation, as it is the result of a process which is global to the image (one cannot construct the watershed segmentation only from local information).

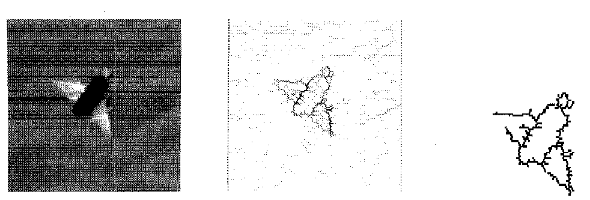


Fig. 3. A hysteresis thresholding on the watershed, valuated by the modulus of the gradient, yields non closed contours and small barbs.

In mathematical morphology, we choose the contours by smoothing the modulus of the image gradient $\|\nabla f\|$ with respect to various criteria. Indeed, we do not work directly on the watershed, but come back to the original information contained in the gradient image. For example, we can suppress some minima while preserving the position of the watershed (this is done by using some **markers**), or we can choose the contour by giving them a value which relies on the values of the gradient and on the watershed (this is what we call **hierarchical segmentation**).

All these morphological methods rely on an arbitrary flooding process, and the theory of morphological simplification of images can be deduced from a powerful tool: the geodesic reconstruction.

3.1 Geodesic Reconstruction

The geodesic reconstruction was originally developed by Beucher [2]. Let M and N be two closed sets of the plane. We denote by d_M the geodesic distance in M, i.e., the lower bound of the lengths of the paths in M linking a to b in M.

Definition 3.1. *We call* ***geodesic dilation of infinite size*** *of N in M, or* ***geodesic reconstruction*** *of N in M, the (arc) d_M-connected components of M which contain at least one point of N. These components correspond to the points at finite d_M-distance of N. We denote this transformation by $D_M^\infty(N)$. The set N is called the* ***marker set.***

On the lattice, the notation $D_M^\infty(N)$ is formally justified by the formula

$$D_M^\infty(N) = \left[(N \oplus B) \cap M\right]^\infty \quad (3)$$

where $\oplus$ denotes the morphological dilation and B the unit ball of the lattice.

With such a definition, the geodesic reconstruction is a binary transform. The simplest way to extend a binary transform to a gray-tone transform is to describe a function f with the help of its lower threshold $[f]\lambda := \{a \mid f(a) \geq \lambda\}$.

Consider two functions $f \leq g$, the **geodesic reconstruction** of f in g is defined through its lower thresholds:

$$\left[D_g^\infty(f)\right]_\lambda := D_{[g]_\lambda}^\infty\left([f]_\lambda\right) \quad (4)$$

As in the binary case, we call $D_g^\infty(f)$ **(gray-scale) reconstruction of** f **under** g **by dilation**. The function f is often called the **marker function** or **markers**. We focus our attention on the dual of the geodesic dilation of infinite size, the geodesic erosion of infinite size, or **(gray-scale) reconstruction of** f **over** g **by erosion**. The geodesic erosion behaves in a complementary way, that is to say we have to write the functions using their upper threshold $[f]^\lambda := \{x, f(x) \leq \lambda\}$. We then have the following formula:

$$\left[E_g^\infty(f)\right]^\lambda := D_{[g]^\lambda}^\infty\left([f]^\lambda\right) \quad (5)$$

Here, the marker function is g. Using this formula, we have derived a new algorithm for the geodesic erosion, which proceeds by flooding. It is described in the last section devoted to algorithms. This formula is fundamental for the understanding of the watershed. Eroding f over g is done by flooding progressively the catchment basins of g while the meeting with f is not achieved (for more details, see Section 6.1.3). Thus we understand why the smoothing by geodesic erosion is so helpful for the segmentation by watershed: by calculating a geodesic reconstruction, we are implicitly constructing a watershed.

Before explaining how to use geodesic reconstruction, we need to state a theorem which has never been explicitly written, but which was implicitly used by all the authors.

THEOREM 3.2. *Let f and g be two functions from $\mathbb{R}^n$ to $\mathbb{R}$, $f \geq g$. Each regional minimum of the geodesic erosion $E_g^\infty(f)$ contains at least one regional minimum of f.*

That is to say, if $f \geq g$, the geodesic erosion $E_g^\infty(f)$ can only suppress or merge regional minima of f. As the main problem in watershed segmentation is to suppress spurious minima, we understand why geodesic erosion can be so helpful.

3.2 The Technique of Markers: A Geodesic Reconstruction by Erosion of the Marker Function Over the Gradient

The oversegmentation produced by the coarse application of the watershed is due to the fact that each regional minimum gives rise to a catchment basin. However, all the catchment basins do not have the same importance. There are important ones, but some of them are induced by the noise, others are minor structures in the image.

The first type of information one can extract is of a geometrical nature. Suppose we know a connected set of points belonging to an object (or a connected set for each object if there is more than one object to segment), and a set of points belonging to the background. We call these connected components **markers**. If we could modify the image on which to compute the watershed by imposing these sets as regional minima, we then obtain a watershed which has a loop around each object, as each catchment basin represents either the background or one unique object.

This is how to impose some regional minima M on an image f. We construct the image:

$$g(x) = \begin{cases} +\infty & \text{if } x \notin M \\ 0 & \text{if } x \in M \end{cases} \quad (6)$$

This image has the regional minima we want. To keep the information of the original image and to put the watershed on the plateau at height ∞ of g, we geodesically erode g on function $f \wedge g$, i.e., we compute $E_{f \wedge g}^\infty(g)$. Here, we denote $f \wedge g(a) := \min(f(a), g(a))$. This image has the same minima as g, for the geodesic erosion does not add any new minima (Theorem 3.2). Furthermore, all the pixels which are sufficiently high and not in an unselected regional minimum are the same. We can then apply the watershed algorithm on this new image. Fig. 4 illustrates this method and Fig. 5 shows the results on the image of a cook-stove. Note that we can choose any contrast images: here we have replaced the usual gradient modulus by a top-hat transformation $f - f_B$, where $f_B = (f \ominus B) \oplus B$.

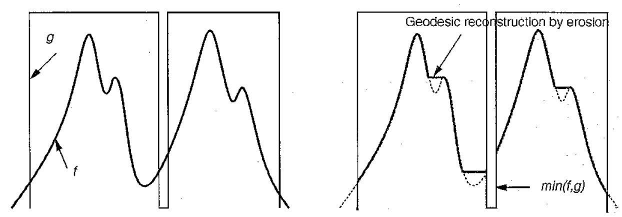


Fig. 4. Illustration of the way to impose regional minima of g on the image f.

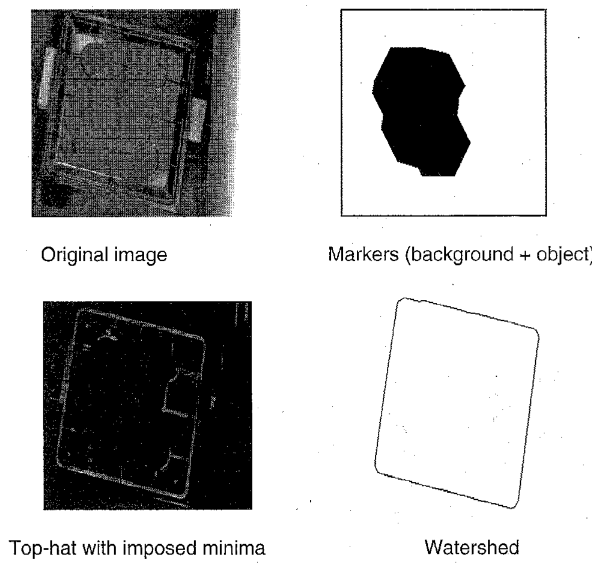


Fig. 5. Constrained watershed: markers are imposed as regional minima of the top-hat transformation.

We did not explain how to choose the markers. It is, in general, the most complex part. The technique of constrained watershed allows us to look for the contour of the objects with less exactitude and guarantees the number of contours found: one around each marked object. All the difficulty lies in determining the markers, i.e., to a rough localization of the objects.

In brief:

Segmentation by constrained watershed

1) Find the markers, i.e., one connected component for each object and one connected component for the background.
2) Compute the image on which the watershed will be constructed (usually a contrast image like the modulus of the gradient).
3) Impose the minima by gray-scale geodesic reconstruction.
4) Compute the watershed.

3.3 The Technique of Minima Dynamics: Geodesic Reconstruction by Erosion of f_t over f

The **dynamics** of a regional minimum is a contrast criterion. Recall that a regional minimum is a connected set from which it is impossible to reach a point with a lower height without climbing. The minimal height of this climbing is the valuation of the contrast of the regional minimum.

DEFINITION 3.3. *Let M be a regional minimum of the function f. The* ***dynamics*** [9], [10] *of M is the number*

$$\min\left\{\max_{s\in[0,1]}\{f(\gamma(s)) - f(\gamma(0))\} \mid \gamma : [0, 1] \to \mathbb{R}^2,\ f(\gamma(1)) < f(\gamma(0)), \gamma(0) \in M\right\}$$

where γ *is a path linking two points.*

One can notice that the dynamics is not defined for the global minimum of the image. In practice, however, the image f has a compact domain of definition, and we can always suppose the global minimum is on the boundary of this image, which allows the valuation of the global minimum inside the domain of definition of f.

The concept of dynamics is illustrated in Fig. 6. It can be used to find relevant markers: the minima with a great dynamics. Let us notice that in practice we do not impose these minima by geodesic reconstruction of a marked function. On the contrary, we suppress the regional minima of f with a dynamics lower than a given contrast value t. The standard algorithm to do this operation is to compute the geodesic reconstruction by erosion $E_f^{\infty}(f_t)$ of f_t over f where $f_t(a) = f(a) + t$ (Fig. 7).

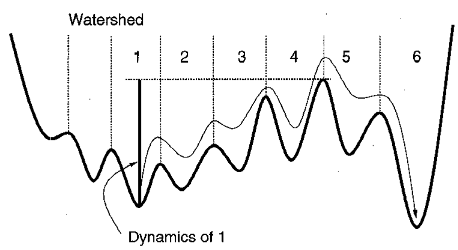


Fig. 6. Illustration of the concept of dynamics.

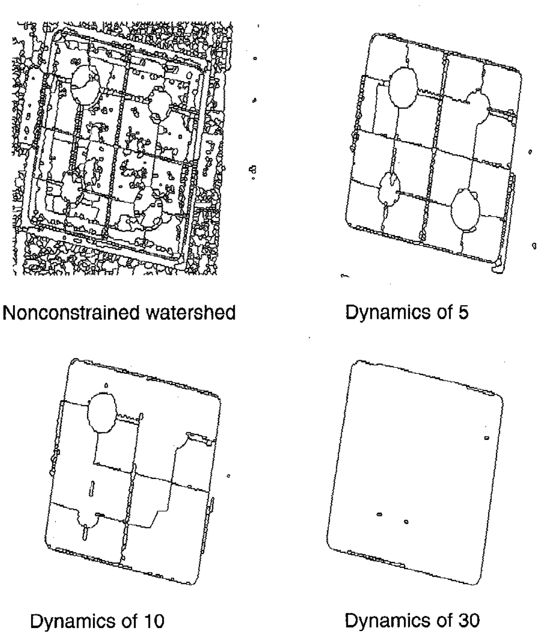


Fig. 7. Watershed constrained by a contrast criterion: the dynamics.

Let us notice that the size and location of the minima with a given dynamics is of little direct use: their catchment basins do not reflect the final segmentation with this level of dynamics (see, later, Fig. 10c). In what follows, we shall explain how to use all the information contained in the map of basins valuated by their dynamics.

The following section is the heart of the paper, in which we create a saliency map from the watershed segmentation.

4 HIERARCHICAL SEGMENTATION

Until now, our aim has been to prevent the oversegmentation by choosing markers and, using homotopy modification, to produce as many catchment basins as there are objects in the image. In this section, we present the notion of **hierarchical segmentation**, originally developed by Beucher [2], which, rather than preventing the oversegmentation, computes the importance of the contours with respect to given criteria.

Let us first mathematically define what we mean by a hierarchy.

DEFINITION 4.1. *Let* $\mathcal{P}_{h_i}$ *be a sequence of partitions of the plane. The family* $(\mathcal{P}_{h_i})_i$ *is called a* ***hierarchy*** *if* $h_i \geq h_j$ *implies* $\mathcal{P}_{h_i} \supseteq \mathcal{P}_{h_j}$*, i.e., any region of partition* $\mathcal{P}_{h_i}$ *is a disjoint union of regions of partition* $\mathcal{P}_{h_j}$.

Every hierarchy can be assigned a **saliency map**, by valuating each point of the plane by the highest value h such that it appears in the boundaries of partition $\mathcal{P}_h$. If we interpret these partitions as segmentations, we have a nice way of assigning importance to the contours. The problem is to obtain such a family of segmentations.

4.1 Beucher's Hierarchical Segmentation: The Waterfall Algorithm

In segmentation with the help of markers, the final result strongly depends on the first stage of marker determination. But we point out that marker determination is not an easy process. Images are often noisy, and the objects we want to detect are often complex and varied in shape, size or intensity. Now, when we look at the result of a watershed segmentation, we notice that a lot of apparently homogeneous regions are shattered into small pieces. A natural idea is then to try to merge these regions. Mathematical morphology suggests a solution to make this fusion, **hierarchical segmentation** which was introduced in this context by Beucher [2] and Beucher and Meyer [4].

This fusion is done by automatically selecting some markers, using a procedure called the **waterfall algorithm** which relies on geodesic reconstruction by erosion. Let us build a new function g by setting $g(x) = f(x)$ if x belongs to the watershed and $g(x) = +\infty$ if not. This function g is obviously greater than f. Let us now reconstruct f over g. It is easy to see that the minima of the resulting image (Fig. 8) are significant markers of the original image.

Some remarks should be made here. First of all, even if this procedure allows the construction of a hierarchy by repeating itself until convergence, it does not allow a valuation of the contours thus obtained: The convergence is

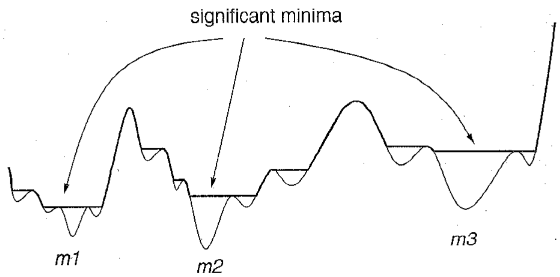


Fig. 8. Reconstruction from watershed lines and detection of the significant markers.

usually very fast, and only a few levels of hierarchy are present in the result (typically half a dozen). Second, even if we value the hierarchy in this way, the final valuation result does not rely on some gradient information: it is only a way of giving a partial order relation on the various minima. A example of such a valuation is given in Fig. 10b.

4.2 Hierarchical Segmentation Using Dynamics

The dynamics of minima notion presented above allows the creation of another hierarchical concept which, by relying on the minima dynamics concept, gives birth to the new concept of dynamics of contour. Let us consider an image f. If we suppose that two different minima and two different saddle points are not on the same gray level (which does not pose any problem in practice), the geodesic reconstruction does not move the contour obtained by watershed. It can only suppress some contours. It is then sufficient to valuate each arc of the watershed with the maximal value of t for which the arc belongs to the watershed of $E_f^\infty(f_t)$. It is easy to see that this depends only on the lowest saddle point on the arcs which separate the two basins. Let a be the (saddle) point of lower altitude on these arcs, we define

$$Bas(a) := \{b \mid \exists \gamma, \gamma(0) = a, \gamma(1) = b, f(\gamma(s)) < f(a)\ \forall s \in]0,1]\} \quad (7)$$

The set $Bas(a)$ is a topological open set, and can be divided in several open connected components B_i ($Bas(a) = \cup_i B_i$). We set

$$dyn(a) := \min_i \max_{a_i \in B_i} \{f(a) - f(a_i)\} \quad (8)$$

We then valuate arcs which separate two basins by the number $dyn(a)$, which we call **dynamics of contour**. The saliency map dyn which associates at each point a its contour dynamics is then given by the formula

$$dyn(a) = \int_0^{+\infty} \chi_{WS(E_f^\infty(f_t))}(a)\,dt \quad (9)$$

where $\chi_{WS(f)}(a)$ is the value at point a of characteristic function[1] of the watershed of f.

This valuation is much more natural than the valuation obtained by the waterfall algorithm, for it relies on the gradient information.

In the last part of this paper, we give an algorithm which directly computes the watershed with this valuation.

1. Which is equal to 1 if a belongs to the watershed and to 0 if a does not belong to the watershed.

It is worth noting that contrary to the noise sensitivity of the dynamics of a basin, the dynamics of a contour is much more stable (yet the contour itself is sensitive to noise). This is illustrated in Fig. 9.

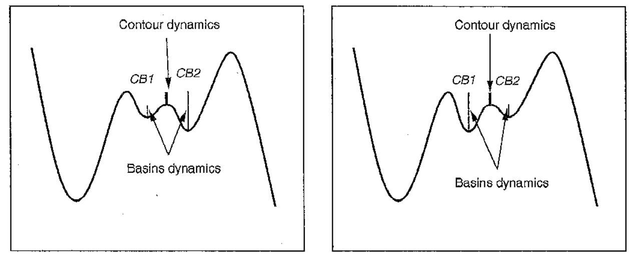
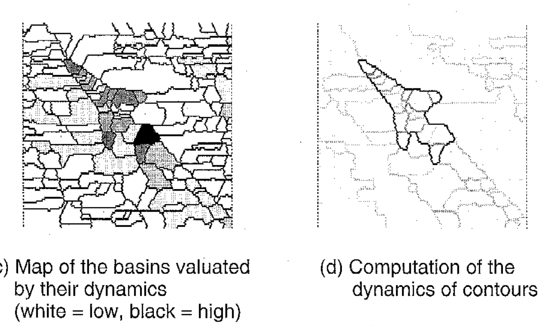


Fig. 9. Between right and left figure, basins CB_1 and CB_2 have exchanged their dynamics. But the dynamics of the contour which separates these basins remains unchanged.

Fig. 10 shows the difference between a simple computation of dynamics (Fig. 10c) and an application of the contour valuation algorithm. The basins of high dynamics do not reflect the extension and shape of the region which could be obtained by keeping only the regional minima of high dynamics. So, the dynamics of basins represent only an intermediate result. Contrary to this, let us notice that the result of our algorithm (Fig. 10d.) gives all the contour information we can extract from the gradient image, that is to say, a threshold of the result image at a given level will give the segmentation which will be obtained by a geodesic reconstruction of the same level. The only way to obtain other contours is to add exterior information, either by the use of markers, either by using other contour extractors (like Canny's one) which we combine to the watershed by use of a watershed algorithm with anchor points [16], [15], [14].



(a) Original image

(b) Waterfall algorithm

(c) Map of the basins valuated by their dynamics (white = low, black = high)

(d) Computation of the dynamics of contours

Fig. 10. Difference between an application of the waterfall algorithm, a computation of dynamics of basins and a computation of the dynamics of contours.

5 Application to Shape Recognition

In this section, we discuss the interest of our new segmentation process. This section is based first on mathematical criteria used in contour detection, then on the robustness of the arc saliency with respect to noise and finally on its usefulness in image segmentation, especially for the problem of guessing "what are the n most important objects in an image."

5.1 Mathematical Arguments

Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a smooth function, and set $f_n(a) := \frac{E(2^n f(a))}{2^n f(a)}$, where $E(x)$ stands for the integer part of x. It has been shown [14], [16] that under adequate hypothesis, the limit of watershed of the f_n is a subset of the maximal integral lines of the gradient of f (lines of steepest slope on f) linking some particular critical points of f (where $\nabla f(a) = 0$), typically a subset of saddle points and of maxima of f. We can apply this result to show why the watershed is a good edge detector.

Let g be the modulus of the gradient of an image f: $g(a) := \|\nabla f(a)\|$. One can say that an edge is a path where the change in the intensity of f is maximum in the direction normal to this path. As the intensity is computed by the modulus of the gradient, we can write

$$\frac{d}{dt} g(a + t\,n) = \langle H_f(a)\nabla f, n \rangle = 0 \qquad (10)$$

where n is the normal to the path at point a, and where $H_f(a)$ is the Hessian (the matrix of the second derivatives) of f at a. This equation gives an implicitly differential equation for the edge path γ:

$$\dot{\gamma} = H_f \nabla f. \qquad (11)$$

This equation is not sufficient to characterize edges, because on any point in the plane where $\nabla f \neq 0$, there exists a path γ satisfying (11). The union of all the paths which are solutions to this differential equation covers the whole domain of f. It has been proved [16], [14] that the watershed chooses the paths by imposing boundary conditions.

The boundary conditions imposed by the watershed are such that it is possible to join the end points of the path $\dot{\gamma} = H_f \nabla f$ to two different minima by two different always descending paths. So the watershed produces closed contours, and finds exactly multiple points (intersection of contours) which are of great importance in image analysis.

On the other hand, classical edge detectors like Canny's[2] [5] solve the problem by **estimating** the normal n from the gradient direction, i.e., by setting $n = \nabla f$ (which is true on a step edge, but not true on more complicated shapes and especially at locations where many contours meet).

Finally, note that the watershed does not need the complete gradient of f (∇f), but only its modulus ($\| \nabla f \|$). This feature is very important near contour junctions, where the direction of the gradient computed with Canny's method is unreliable, yielding very few triple points in its contour images. Some complex modifications have to be made in order to detect the corner points correctly [7], [8]. On the contrary, the watershed, using only the modulus of the gradient, (which is reliable) gives triple points, accurately positioned. This feature has been successfully applied for corner detection in [17].

2. Canny's detector, or more exactly the extrema of the gradient in the direction of the gradient, finds the zero crossings of $Q(f) = \langle H_f \nabla f, \nabla f \rangle$.

5.2 Noise Robustness

One kind of performance evaluation for a segmentation algorithm is its small dependence to noise. Fig. 11 shows a synthetic image with its watershed computed directly on it. Due to the noise free image, all the watershed lines have the same saliency and the four square shaped catchment basins are very large. When adding a small negative Poisson noise, we observe the birth of many small catchment basins created by one pixel large regional minima corresponding to negative noise. The noise has been chosen negative, because at each modified point, it creates a small regional minimum and, consequently, a new catchment basin. However, the new added watershed lines have a very low saliency, whereas the most contrasted watershed lines are very little displaced. The small displacements we observe correspond to noise pixels falling exactly on the watershed line. Also, the junction point at the center of the image, where the four major contours meet, is always preserved. Increasing the noise values does not add new basins, but the saliency of their contours increases. Note however that the most salient

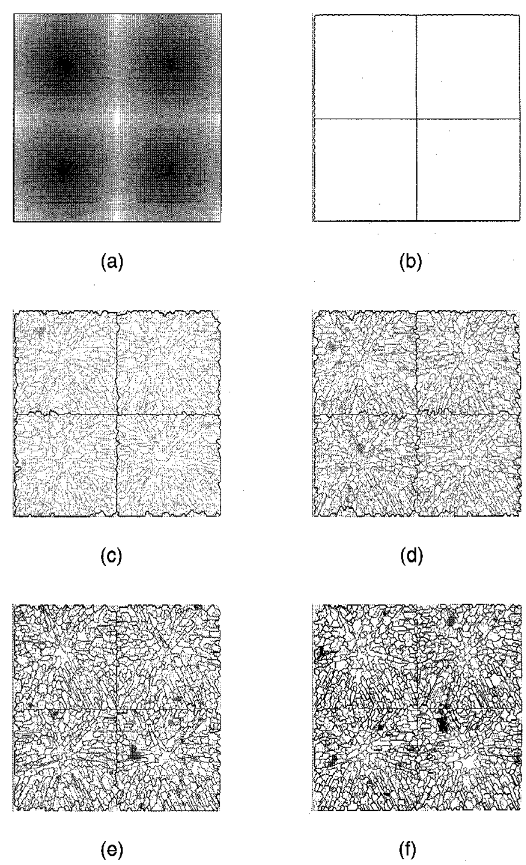


Fig. 11. Robustness of the dynamics of contours with respect to noise. (a) original noise free image, (b) its watershed, (c) to (f) increasing negative Poisson noise.

contours, namely the desired contours, are still very little affected. Finally, when the noise completely overwhelms the signal (noise values similar to the signal values), the most salient contours no longer really exist. In this case, optimal linear smoothing techniques should be applied prior to watershed extraction.

5.3 Real Image Segmentation

Let us now illustrate the concept of hierarchy on real images. Fig. 12 shows some indoor scene, where some of the structures are very contrasted, whereas others are much less so. These structures can be recognized on the saliency map. The very low salient contours correspond to noise. Note however that some structures like the upper part of the desk exhibit a lower saliency than expected. This is due to the criterion implemented by the watershed: the saliency is governed by the altitude of the saddle point with respect to the two neighboring catchment basins. The altitude of this saddle point corresponds to a minimum on the contour line. So, as soon as some parts of the contour are less contrasted, even for a few pixels (here at the left of the lamp), the lower contrast value is propagated along the whole object contour. In particular, objects which exhibit a low contrasted side are surrounded by a watershed line of low saliency. This is due to the fact that the watershed always closes the contours, yielding a region based segmentation.

Another example is presented in Fig. 13.

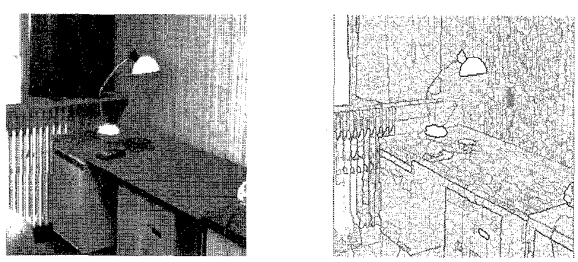


Fig. 12. Indoor scene and the saliency map obtained by the watershed of the modulus of its gradient.

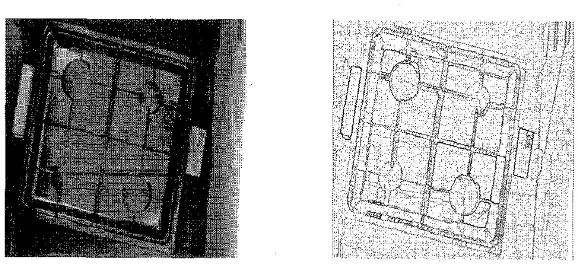
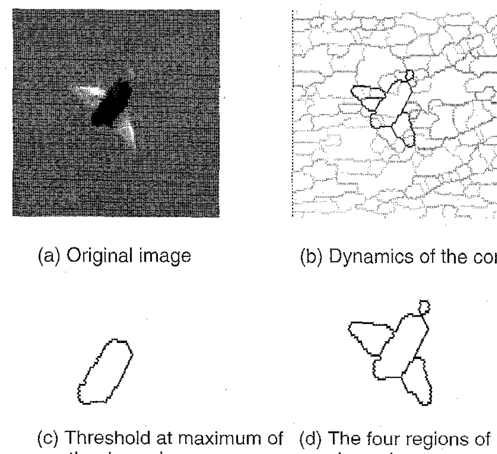


Fig. 13. Another example of the saliency map computed on an image of a cook stove.

5.4 Finding n Objects

In this section, we give an example showing how to use the dynamic hierarchical segmentation algorithm for shape recognition.

Fig. 14 is a snapshot of an airplane. For military systems, one of the key problems is airplane identification and attitude estimation. With this in mind, we may construct a



(a) Original image (b) Dynamics of the contours

(c) Threshold at maximum of the dynamics (d) The four regions of higher dynamics

Fig. 14. Some thresholds of the contour dynamics image.

three-dimensional model of the plane, which will be compared to a database of three-dimensional models. More simply, we may compare only the contour of the airplane to a contour database, by extracting a small number of parameters which hopefully reflect the relevant features of the shape [18], [19].

So, the first stage consists of finding the airplane contours. Here the use of markers is very difficult: To obtain the whole plane correctly, including the wings, the marker should have the shape of the plane and, as a marker, it should be included in the plane. In the same way, the use of the highest dynamics will only extract one region in some cases, and we will obtain only the airplane body (Fig. 14c). We propose the use of the dynamic hierarchical segmentation to extract the right contours.

The image of the contour dynamics (Fig. 14b) clearly shows that all the interesting contours are present in the watershed image. One has to introduce additional information for extracting these right contours. For instance, if we choose the object surface, the image of the contour dynamics will be thresholded at a height corresponding to n catchment basins of the expected surface. One can also consider applying criteria for region growing [13], the initial step being a threshold of the image of the contour dynamics.

However, finding the airplane contour directly is futile. On the other hand, one can guarantee that, in a small number of hypotheses, we will find this right segmentation. This is done by successively merging regions to construct the airplane shape. The algorithm is a good procedure for giving these hypotheses (four regions, Fig. 14d) and one can identify the airplane.

In the same way, in the context of interactive segmentation, the hierarchical saliency map can be computed at once. Then, the adjustment of the unique threshold, done manually, allows a human operator to explore the various possible segmentations.

6 ALGORITHMS

6.1 Algorithms for Geodesic Reconstruction

6.1.1 Sequential Reconstruction

There exists a sequential algorithm [11], [23], [24] for geodesic reconstruction which works by propagating information downward in a raster scanning and then upward in an antiraster scanning. These raster and antiraster scannings have to be iterated until stability is reached (usually no more than ten complete image scannings). It is then very fast.

6.1.2 Beucher-Meyer Algorithm

The principle [4] is simple. The regional minima of $g > f$ are given as input. Starting from these minima, g is eroded progressively, while staying above f. The implementation of this algorithm is easy with an ordered queue [4] or with a heapsort algorithm [1] (these two algorithms create a queue which stays ordered when a new element is introduced): The pixels are examined by increasing gray level of g. Graphically speaking, the function g acts as a film which contracts on a parcel which is the function f. The algorithm is optimal in the sense that each point is processed only once.

6.1.3 New Algorithm

We now present a new algorithm for geodesic reconstruction which "dilates" f under g, and proceeds by flooding. It is optimal in the same way as the Beucher-Meyer algorithm presented above: Each point is processed only once. But more important, it does not need the regional minima as an input. The algorithm computes the regional minima *during* the flooding, as in the classical watershed algorithm [25].

The basic idea is to use the flooding principle developed by Vincent [23]. This principle is adapted to the geodesic reconstruction of f under g, which corresponds to the geodesic erosion of g with respect to f.

In fact, we can reconstruct f by flooding the catchment basins of f until they overflow, or until we meet g. When we flood the catchment basins of f, two cases can appear.

- Either the minimum of g on the basin is equal to the height of flooding we have reached. In this case, we stop flooding this basin (basin CB_3 in Fig. 15a).
- Either we have filled the basin until one of its saddle points (contact point between two basins) is reached. This case is divided into two branches:

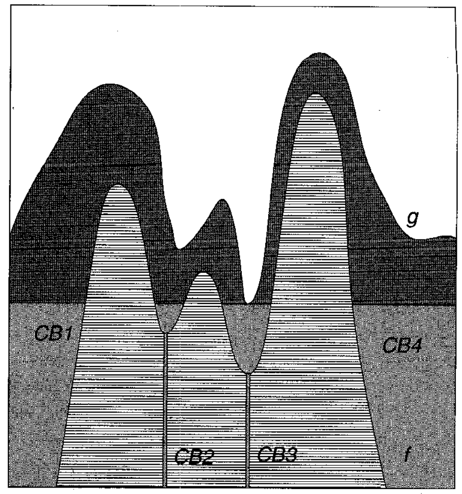
  i) either the other basin has already stopped being flooded. We then stop flooding the considered basin at the height of the flooding (basin CB_2 in Fig. 15b);

  ii) either the other basin has not yet been stopped. We merge the two basins.

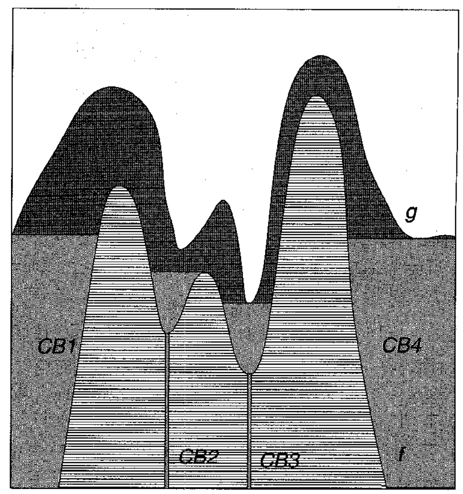
The watershed by flooding algorithm is easy to adapt to this procedure which is fundamental for the computation of constrained watersheds and for the computation of regional maxima under constraints [22]. The advantage of this algorithm with respect to the one of Beucher-Meyer is that the new algorithm computes the regional minima directly.

6.2 Algorithm of Contour Valuation

We now give an algorithm which computes the contour valuation. During a first stage, we use Grimaud's algorithm [9] to compute the watershed, the catchment basins and



(a) At time t, the catchment basin CB_3 stops to be flooded.

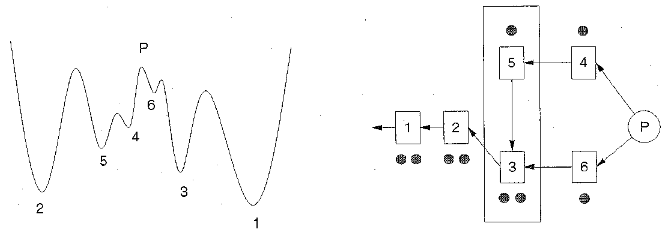


(b) Final stage of the reconstruction.

Fig. 15. Geodesic reconstruction of f under g.

their dynamics. We then compute, for each point of the watershed, a valuation by doing a kind of "gradient descent" on the dynamics value of the catchment basins. It is difficult to compute the valuation of the point during the watershed construction, because when a point of the watershed is created, we do not dispose of the complete list of its neighbors, and in particular triple points can end with a wrong value. In fact, only one point on each arc of the watershed is of interest: the saddle point.

Let us briefly recall Grimaud's dynamics algorithm which is based on Vincent's watershed algorithm. It consists in flooding from the minima, level by level, until water from one minimum meets water from another minimum. The meeting point between two basins is a saddle point, and is the point where we can compute the dynamics of one of the two basins: the basin with the lowest minimum floods the other one, and the dynamics of the basin with the highest minimum is equal to the gray-value of the saddle point minus the gray-value of the minima.



(a) Example of function (b) Flooding list

Fig. 16. Determination of the basin which valuates the contour at which the pixel P belongs.

In our algorithm, the arc of the watershed containing this saddle point is valued by this dynamics. Let us notice that, in Grimaud's algorithm, an arbitrary choice has to be made if the two meeting basins have minima with the same gray-value, but this choice *does not change the value of the arc*. This is why we think that the dynamics of an arc is a better notion than the dynamics of a catchment basin.

We now show how we can value the watershed arc. The best way to explain the algorithm is to look at an example (Fig. 16). Suppose we have applied Grimaud's algorithm on the function of Fig. 16a. We then have a list of watershed pixels, and a list of catchment basins. During the flooding process, we construct a list of catchment basins. Each catchment basin keeps in memory:

- Its dynamics,
- A list of pointers on the catchment basins which have flooded it.

Let's have a look at a pixel p of the watershed (Fig. 16a). This pixel is the connection point between two (or more, three at most on the hexagonal grid) catchment basins. In our example, they are catchment basins 6 and 4.

But basin 6 has been flooded by basin 5, and basin 5 has been flooded by basin 3, which itself has been flooded by basin 2. The basin which has flooded all the other basins is the one with the lowest minimum, that is to say basin 1. In the other way, basin 6 has been flooded by basin 3. This flooding list is represented in Fig. 16b.

The pixel p is in fact flooded when basins 3 and 5 meet and the dynamics of the contour can be computed. Another way to choose the dynamics of the pixel p is to notice that pixel p belongs to the interior of basin 3. So we can say that the dynamics of pixel p is the highest dynamics of the basins which precede basin 3 in the flooding list. So, all the difficulty is to mark out basin 3. We propose a simple procedure for doing that operation. It consists of running through the whole flooding list issued from the pixel p, and in marking the basin by which we pass. The first basin which has been passed over more than once contains the pixel p in its interior, and is basin 3 in our example. So the dynamics of the pixel p is the highest dynamics between the dynamics of basins 5 and 6 (which precede basin 3 in the flooding list), that is to say the dynamics of basin 5.

One could verify that the watershed arc which contains pixel p disappears in a geodesic reconstruction of size equal to the dynamics of basin 5.

7 Conclusion

Image segmentation is not a goal in itself. It should be done by keeping in mind the real purpose of the image processing. We have given an algorithm which is useful for interactive dynamics thresholding. The concept of dynamics of contours allows the valuation of the contours produced by the watershed. This resulting segmentation is *identical* to the one obtained by watershedding the original gradient image after a filtering by reconstruction with the same contrast value. The advantage of the proposed algorithm is this : the dynamic segmentation can be obtained for different contrast values by simply thresholding the valuated watershed image. For instance it enables us to answer the question "which are the n most contrasted objects?" by a simple thresholding of the image (a problem which can be solved in a more complicated way by examining the histogram of dynamics of the image minima).

An example of image processing is **identification**. It is almost impossible to obtain the right segmentation for identification using only segmentation tools based on contrast. But we can guarantee that, in a small number of hypotheses, we will find this right segmentation, which can be extracted by artificial intelligence tools (using for instance a database). We expect our algorithm to be a good procedure for giving these hypotheses.

8 Notations

a, b points in the $\mathbb{R}^2$ plane
f, g images, i.e., functions from $\mathbb{R}^2$ to $\mathbb{R}$
∇f gradient of f, i.e., vector of first derivatives
H_f Hessian of f, *i.e.*, matrix of second derivatives
$\langle . , . \rangle$ inner product
$f \wedge g$ pointwise minimum of f and g
$f \vee g$ pointwise maximum of f and g
$[f]^h$ the upper threshold of f at level h, i.e., $\{a \in \mathbb{R}^n \mid f(a) \leq h\}$
$[f]_h$ the lower threshold of f at level h, i.e., $\{a \in \mathbb{R}^n \mid f(a) \geq h\}$
$f \oplus B$ dilation, i.e., $\max\{f(y) \mid y \in B_x\}$
$f \ominus B$ erosion, i.e., $\min\{f(y) \mid y \in B_x\}$
$D_g^\infty(f)$ geodesic reconstruction of f into g by dilation ($f \leq g$)
$E_g^\infty(f)$ geodesic reconstruction of f over g by erosion ($f \geq g$)
$d_M(a, b)$ geodesic distance in M (Definition 2.1)
$SKIZ_A(B)$ skeleton by influence zones of B according to the geodesic distance in A (Definition 2.2)
$IZ_A(B)$ influence zones of B according to the geodesic distance in A (Definition 2.2)
Reg_Min_h (f) regional minima of f *i.e.*, connected plateau from which it is impossible to reach a point of lower gray level by an always descending path

REFERENCES

[1] A. Aho, J. Hopcroft, and J. Ullman, *Data Structures and Algorithms.* Reading, Mass: Addison-Wesley, 1983.

[2] S. Beucher, "Segmentation d'images et morphologie mathématique," thèse, École Nationale Supérieure des Mines de Paris, June 1990.

[3] S. Beucher and C. Lantuéjoul, "Use of Watersheds in Contour Detection," *Proc. Int'l Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France, Sept. 17-21, 1979.

[4] S. Beucher and F. Meyer, "The Morphological Approach to Segmentation: The Watershed Transformation," *Mathematical Morphology in Image Processing*, E.R. Dougherty, ed., Optical engineering, pp. 433-482. New York, Basel, Hong Kong: Marcel Dekker, 1993.

[5] J.F. Canny, "A Computational Approach to Edge Detection," *Readings in Computer Vision: Issues, Problems, Principles and Paradigms*, M.A. Fischler and O. Firschein, eds., pp. 184-203. Morgan Kaufmann, 1986.

[6] H. Digabel and C. Lantuéjoul, "Iterative Algorithms," *Proc. Second European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, J.-L. Chermant, ed., pp. 85-99. Stuttgart, Germany: Riederer Verlag, 1978.

[7] G. Giraudon and R. Deriche, "Accurate Corner Detection: An Analytical Study," Technical Report 1420, INRIA, Apr. 1991.

[8] G. Giraudon and R. Deriche, "On Corner and Vertex Detection," Technical Report 1439, INRIA, June 1991.

[9] M. Grimau, "La Géodésie Numérique en Morphologie Mathématique: Application à la Détection Automatique de Microcalcifications en Mammographie Numérique," thesis, École des Mines de Paris, Dec. 1991.

[10] M. Grimaud, "A New Measure of Contrast: Dynamics," *SPIE Vol. 1769, Image Algebra and Morphological Processing III*, pp. 292-305, San Diego, July 1992.

[11] B. Läy, "Recursive Algorithms in Mathematical Morphology," *Acta Stereologica*, pp. 691-696, Caen, France, 1987, vol. 6/III, *Proc. Seventh Int'l Congress Stereology.*

[12] F. Meyer, "Topographic Distance and Watershed Lines," *Signal Processing*, special issue on mathematical morphology, vol. 38, no. 1, pp. 113-126, July 1996.

[13] O. Monga, "An Optimal Region Growing Algorithm for Image Segmentation," *Int'l J. Pattern Recognition and Artificial Intelligence*, vol. 3, no. 4, Dec. 1987.

[14] L. Najman, "Morphologie Mathématique: de la Segmentation d'Images à l'Analyse Multivoque," thèse de doctorat, Université Paris-Dauphine, Paris, France, Apr. 1994.

[15] L. Najman and M. Schmitt, "La Ligne de Partage des eaux: applications d'une approche continue," *Revue Technique Thomson*, vol. 25, no. 2, pp. 261-280, Mar. 1993.

[16] L. Najman and M. Schmitt, "Watershed of a Continuous Function," *Signal Processing*, special issue on mathematical morphology, vol. 38, no. 1, pp. 99-112, July 1994.

[17] L. Najman and R. Vaillant, "Topological and Geometrical Corners by Watershed," *CAIP '95 Proc.*, Hlavác and Sára, eds., LNCS 970. Springer-Verlag, 1995.

[18] L. Najman, R. Vaillant, and É. Pernot, "From Face Sideviews to Identification," *Revue Technique Thomson*, vol. 24, no. 4, pp. 1,037-1,054, Dec. 1992.

[19] L. Najman, R. Vaillant, and É. Pernot, "From Face Sideviews to Identification," *Image Processing: Theory and Applications, An Int'l Conf.*, G. Vemazza, ed., pp. 299-302, San Remo, Italy, June 1993.

[20] M. Schmitt, "Geodesic Arcs in Non-Euclidean Metrics: Application to the Propagation Function," *Revue d'Intelligence Artificielle*, vol. 3, no. 2, pp. 43-76, 1989.

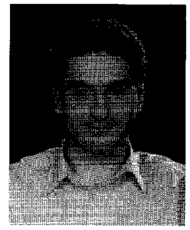
[21] M. Schmitt and J. Matioli, *Morphologie Mathématique.* Logique—Mathématiques—Informatique. Masson, Dec. 1993.

[22] M. Schmitt and L. Vincent, *Morphological Image Analysis: A Practical and Algorithmic Handbook.* Cambridge Univ. Press, to appear in 1996.

[23] L. Vincent, "Algorithmes Morphologiques à Base de Files d'Attente et de Lacets: Extension aux Graphes," thèse, École des Mines, Paris, May 1990.

[24] L. Vincent, "Morphological Algorithms," *Mathematical Mrophology in Image Processing*, E.R. Dougherty, ed., Optical engineering, pp. 255-288. New York,Basel, Hong Kong: Marcel Dekker, 1993.

[25] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, June 1991.



Laurent Najman received his Maîtrise de Mathématiques Appliquées and his Maîtrise d'Ingénierie Mathématique in 1989 from the University of Paris VI, his engineering degree from the Ecole des Mines de Paris in 1991, and the Docteur en Sciences degree from the Paris-Dauphine university. From 1991 to 1994, he was in the Signal Analysis and Pattern Recognition Laboratory at the Central Research Laboratory of Thomson-CSF, where he worked in collaboration with the Viability Team of the CEREMADE at Paris-Dauphine University on some problems of image segmentation using mathematical morphology and set valued analysis. He joined ArSciMed in 1995, a start-up company devoted to the development of particle systems technology, where he is director of research and development. His current research interest is simulation using particle systems, with applications to computer graphics and vision.

Michel Schmitt received his engineer degree from the Ecole Polytechnique in 1982, his engineer of the Corps des Mines degree in 1985. He joined the Mathematical Morpholgical Laboratory of J. Serra at the Ecole des Mines (1985) where he began a thesis on mathematical morphology: theoretical aspects, algorithms, and artificial intelligence. He joined the team of O.D. Faugeras at INRIA in 1985, carrying on his research work, obtained his Docteur en Morphologie Mathématique degree from the Ecole des Mines in 1989 and the Habilitation à diriger des recherches degree in 1991. From 1989 to 1995, he worked at the Central Research Laboratory of Thomson-CSF as head of the Signal Analysis and Pattern Recognition Laboratory. Since 1995, he has been the director of the center for Geostatistics at the Ecole des Mines de Paris. His current research interests are image processing and geostatistics.