ELSEVIER

Signal Processing 38 (1994) 113–125

SIGNAL PROCESSING

Topographic distance and watershed lines

Fernand Meyer*

Centre de Morphologie Mathématique, 35 Rue Saint Honoré, 77305 Fontainebleau, France

Received 5 October 1993; revised 5 December 1993

Abstract

The watershed line is the basic tool for segmenting images in mathematical morphology. A rigorous definition is given in terms of a distance function called topographic distance. If the topographical function is itself a distance function, then the topographical distance becomes identical with the geodesic distance function and the watershed becomes identical with the skeleton by zone of influence. The classical shortest paths algorithms of the graph theory are then revisited in order to derive new watershed algorithms, which are either new or more easy to implement in hardware.

Zusammenfassung

Die Wasserscheide ist das meist benutztze Werkzeug zur Bildsegmentierung, das die Mathematische Morphologie entworfen hat. Eine topographische Distanzfunktion wird eingeführt, die es erlaubt der Wasserscheide eine präzise mathematische Definition zu verleihen. Auf eine normale Distanzfunktion angewandt ist die topographische Distanz mit der geodätischen Distanz gleich, sodass die Wasserscheide dann zum Skeleton bei Influenz Zone wird. Klassische graph-theoretische kürzesten Weg Algorithmen können dann in diesem Rahmen angewandt werden. Es ergibt sich daraus neue Algorithmen zur Bestimmung von Wasserscheiden, die für die einen präziser als die üblichen sind, und die für die anderen leichter in eine Rechner Architektur einsetzbar sind.

Résumé

La ligne de partage des eaux est la pierre angulaire de la segmentation en Morphologie Mathématique. Une définition rigoureuse en est proposée en termes de distance topographique. Sur un relief qui est lui-même une fonction distance, la distance topographique se réduit à la distance géodésique et la ligne de partage des eaux devient le squelette par zone d'influence (SKIZ). Les algorithmes de chemins minimaux sur graphe sont revisités et permettent de proposer des algorithmes nouveaux de construction de la ligne de partage des eaux; certains d'entre eux sont plus précis, d'autres plus faciles à réaliser en hardware.

Key words: Topographic distance; Watershed line; Segmentation; Shortest path algorithms; Mathematical morphology

1. Introduction

The watershed line is the key tool developed within the framework of mathematical morphology for segmenting images. The notion has been introduced by Beucher and Lantuéjoul [3]. The watershed algorithm is used mainly on gradient images. It detects the catchment basins of all minima in the gradient image. The best intuitive presentation is due to Beucher who considered the gradient image

*Tel: (33)1 64 69 47 06, Fax: (33)1 64 69 47 07, E-mail: meyer@cmm.ensmp.fr

SSDI 0165-1684(94)00022-R

as a topographical relief. This relief is flooded; the sources are placed at the regional minima. The level of the flood is uniform over the relief and increases with uniform speed. The moment that the floods filling two distinct catchment basins start to merge, a dam is erected in order to prevent mixing of the floods. The union of all dams constitutes the watershed line. Often these minima are far too numerous and most of them are irrelevant for the desired segmentation; which leads to strong oversegmentations. The solution I proposed in 1982 helps to overcome this problem, as soon as the first approximation of the solution has been found. We call this first approximation markers, since it involves an inside marker for each object that is to be detected, including the background. The segmentation is done in two steps: (a) In the first step, a set of markers, is detected for each object and the background. As noted by Beucher [2], this step constitutes the intelligent side of the segmentation and is highly problem dependent. (b) The second step consists in constructing the watershed of the gradient image, by flooding the gradient relief from a set of sources identical to the markers. After completion of the flooding, one gets a partition of the image where each tile contains one and only one of the markers. A complete presentation of this method with many examples of applications may be found in [9, 2]. However, no intrinsic definition of the watershed line and the catchment basins are given in these references: the only definitions are either intuitive presentations, based on topographical analogies, or algorithmic definitions. An attempt was made in [12] to provide a mathematical definition, with a rather unsatisfactory result, since two distinct distance functions are necessary for defining respectively the SKIZ and the watershed line. We would instead like, that the same distance function yields both the watershed line, when applied to a general grey-tone function, and the SKIZ, when applied to a distance function. Only recently, during the first workshop of mathematical morphology in Barcelona, two papers [8, 11] presented a rigorous definition of the watershed line. The definition in [11] asks for a higher regularity than the definition in [8]. For regular functions, however, they are identical. The present paper is an extension of [8].

The background to this work is as follows. An algorithm for integrating images has been published by Verbeek et al. [13]. We have shown in [7] that the corresponding operator σ is the inverse operator of the half-morphological gradient operator $\partial = (I - \varepsilon)$, where ε represents the elementary erosion (defined more precisely below) and I the identity operator. This means

- $\partial\sigma = I$: first integrating the function and defferentiating the result retains the initial image, outside of the set of points used as the limiting conditions for the integration. This constitutes a solution of the eikonal equation.
- $\sigma\partial = I$: it is possible to recover the initial image from its half-morphological gradient by integration, if one takes as limit conditions one point in each regional minimum. This is based on the following property: two grey-tone functions with the same regional minima and the same half-morphological gradient are identical.

This last relation is the key result of [7]. As we have shown, it allows us to establish a link with the watershed algorithm: during the integration, each point gets its value from one and only one ancestor. If we attach the same label to all points having a common ancestor, we get a mosaic image. Each tile of the mosaic is a catchment basic of the reconstructed image. Their boundaries are the watershed lines of the reconstructed image.

As pointed out by Verbeek, the integration algorithm is in fact an algorithm for computing a weighted distance. One may associate to the digital grid a weighted graph. Some nodes of the graph serve as limit conditions and are assigned a grey-tone value. Then, integrating the function ∂f is the same than computing a path with minimal cost between the limit nodes and all other pixels.

If the cost is the half-morphological gradient, the length of a minimal cost paths is a distance function, refered to here as the topographical distance. The catchment basin of a regional minimum is defined as the set of points which are for the topographical distance, closer to this minimum than to any other. The rigorous definition of the watershed line also leads to the design of new and efficient algorithms: On the one hand more precise algorithms, on the other hand algorithms which are easier to implement in hardware.

2. Definition of the watershed in the continuous space

Most often the watershed is constructed on a gradient image; such images are often noisy and not very smooth. For this reason one has to be careful when defining the watershed for functions in the continuous space: the definition should be valid even if the function is not very regular. For this reason, we will try to remain as general as possible in defining the topographical distance function.

2.1. The topographical variation

We consider now a function f from $\mathbb{R}^n$ into $\mathbb{R}$. Let $\mathrm{supp}(f)$ be its support. Let T be an interval of $\mathbb{R}$ and γ be a continuous function from T into $\mathrm{supp}(f)$. (T, γ) is then a path contained in the support of f. Let $\zeta = (t_1 < t_2 < \cdots < t_n)$ be a finite part of T. For the sake of brevity, we write $\gamma_i = \gamma(t_i)$. We further define the elementary erosion ε_i as the erosion by a disk of radius $|\gamma_{i-1}, \gamma_i|$, the geodesic distance between γ_{i-1} and γ_i.

Definition 1. The topographical variation of the function f along the polygonal line ζ is defined as

$$\mathrm{TV}_\zeta = \sum_i [f(\gamma_i) - \varepsilon_i f(\gamma_i)].$$

Remark 1. Let us consider the special case where the function is a distance function. Such a function has a constant slope equal to λ. Hence the erosion of the function by a disk of radius λ is the same as subtracting the constant 1 from the function (outside the regional minima and their vicinity). The expression in the topographical variation becomes then: $f(\gamma_i) - \varepsilon_i f(\gamma_i) = |\gamma_{i-1}, \gamma_i|$.

Remark 2. In general we will not have the property $(\zeta \subset \zeta') \Rightarrow (\mathrm{TV}_\zeta \leqslant \mathrm{TV}_{\zeta'})$ except for distance functions. For other functions however, as we will see, it holds in the vicinity of the lines of biggest slope. This is sufficient for our purpose.

Definition 2. We call topographical variation of the function f on the path (T, γ), the positive number, finite or infinite, defined by: $\mathrm{TV}_\gamma = \sup TV_\zeta$ (for all finite $\zeta \subset T$).

If $\mathrm{TV}_\gamma < \infty$, one says that f is of finite topographical variation on T.

The topographical distance between two points p and q is then easily defined by considering the set $\Gamma(p, q)$ of all paths between p and q which belong to the support of f:

$$\mathrm{TD}(p,q) = \inf_{\gamma \in \Gamma(p,q)} \mathrm{TV}_\gamma.$$

Proposition 1. *If p and q belong to a line of steepest slope between p and q, $f(q) > f(p)$, then*

$$\mathrm{TD}(p, q) = f(q) - f(p).$$

In all other cases, we have $TD(p, q) > f(q) - f(p)$. Hence the lines of steepest slope are geodesics of the topographic distance function.

Proof. Let γ be a line of steepest slope, and $\zeta = (t_1 < t_2 < \cdots < t_n)$ a set of pixels belonging to γ. Then γ_{i-1} is the lowest pixel in the disk centered at γ_i of radius $|\gamma_{i-1}, \gamma_i|$. For this reason we have $\varepsilon_i f(\gamma_i) = f(\gamma_{i-1})$. Hence $\mathrm{TV}_\zeta = \sum_i [f(\gamma_i) - \varepsilon_i f(\gamma_i)] = \sum [f(\gamma_i) - f(\gamma_{i-1})] = f(q) - f(p)$.

If the path does not follow a line of steepest slope, we will find at least a pixel γ_i for which $\varepsilon_i f(\gamma_i) > f(\gamma_{i-1})$. □

Earlier we observed that, in general, the inequality $(\zeta \subset \zeta') \Rightarrow (\mathrm{TV}_\zeta \leqslant \mathrm{TV}_{\zeta'})$ does not hold. It holds however in the vicinity of the path of steepest slope. This ensures that our definition of the topographical distance is consistent with our goal: we want a distance function yielding the distance $f(q) - f(p)$ for any pixels along a line of steepest slope and yielding much higher values for paths which are very different from paths of steepest slope. Let us consider two polygonal lines ζ and ζ' inscribed in the path γ: $\zeta \subset \zeta'$ and three successive pixels of ζ': $a < b < c$, such that $a \in \zeta$, $c \in \zeta$. The situation is illustrated in Fig. 1.

If the minimum m of the function f inside the disk $(a, |ac|)$ is included in the union of the disks

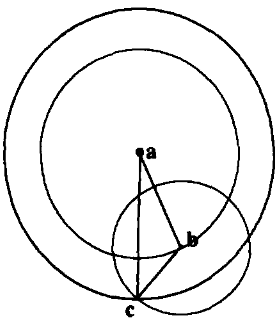

Fig. 1. Triangular inequality on a path.

$(a, |ab|) \cup (b, |bc|)$, then the triangular inequality (1) is satisfied:

$$(f - \varepsilon_{|ac|} f)(a) < (f - \varepsilon_{|ab|} f)(a) + (f - \varepsilon_{|bc|} f)(b). \quad (1)$$

Proof. (a) If the minimum m belongs to the disk $(a, |ab|)$ then $f(a) - \varepsilon_{|ac|} f(a) = (f(a) - \varepsilon_{|ab|}(f(a))$, and the triangular inequality is satisfied.

(b) If the minimum m belongs to the disk $(b, |bc|)$, then

$$f(b) - \varepsilon_{|bc|} f(b) = f(b) - \varepsilon_{|ac|} f(a). \quad (2)$$

On the other hand, it is always true that

$$f(a) - \varepsilon_{|ab|} f(a) > f(a) - f(b). \quad (3)$$

Adding (2) and (3), we get again the inequality (1) □.

If for any triple of points $(t_i < x < t_{i+1})$, with $(t_i, t_{i+1}) \in \zeta^2$ and $x \in \zeta'$, we have the property that the minimum of f inside the disk $(\gamma(t_{i+1}), |\gamma(t_i)\gamma(t_{i+1})|)$ is included in the union of disks $(t_{i+1}, |\gamma(x)\gamma(t_{i+1})|) \cup (x, |\gamma(x)\gamma(t_i)|$. In this case the polygonal lines ζ and ζ' are in the vicinity of a line of steepest descent and the triangular inequality holds for all these points; hence $TV_\zeta \leqslant TV_{\zeta'}$.

Let $(m_i)_{i \in I}$ be the set of regional minima of the function f. We suppose from now on that they have all the same value v. If it is not the case, it will not change the watershed line, if one gives to all of them the value of the deepest.

Definition 3. We call catchment basin $CB(m_i)$ of a regional minimum m_i the set of points $x \in supp(f)$ which are closer to m_i than to any other regional minimum for the topographical distance:

$$\forall j \in I, \quad j \neq i \;\Rightarrow\; TD(x, m_i) < TD(x, m_j).$$

Remark. In the case where the levels of the minima are not the same the definition of the catchment basins $CB(m_i)$ becomes

$$\forall j \in I, \quad j \neq i \;\Rightarrow\; level(m_i) + TD(x, m_i) < level(m_j) + TD(x, m_j).$$

Definition 4. The watershed line of a function f is the set of points of the support of f which do not belong to any catchment basin:

$$Wsh(f) = supp(f) \cap \left[\bigcup_i (CB(m_i)) \right]^c.$$

Remark. If the function f possesses a gradient except at some isolated points, one has the following relation: $f(\gamma_i) - f(\gamma_{i-1}) = |\nabla f(\gamma_i)| * |\gamma_i\gamma_{i-1}| + o(|\gamma_i\gamma_{i-1}|)$, where γ_{i-1} belongs to the line of greatest slope descending from γ_i. The topographical distance function reduces then to

$$TD(p, q) = \inf_{\gamma \in \Gamma(p,q)} \int_\gamma |\nabla f(\gamma(s))| ds. \quad (4)$$

2.2. Application to binary images

One of our requirements for the definition of the watershed line was its applicability to binary images: applied to the distance function to binary sets, the watershed line should yield the SKIZ of these sets. Indeed, the steepest slope of a distance function is everywhere constant and equal to 1. On a distance function the formula (4) is valid and becomes $TD(p, q) = \inf_{\gamma \in \Gamma(p,q)} \int_\gamma ds$, which is the classical formula for the geodesic distance between p and q.

Proposition 2. *For a distance function f, the topographical distance between two points p and q reduces to the geodesic distance within* supp(f) *between these two points.*

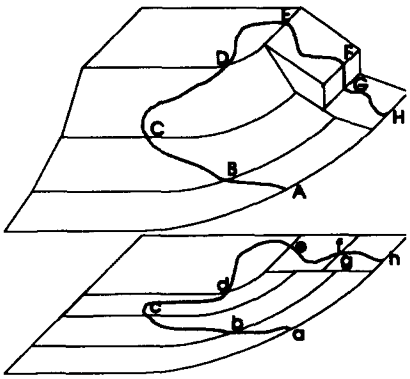

Fig. 2. Relation between the topographic distance along a path and the geodesic distance along the projected path on the support of f.

2.3. *Interpretation of the topographical distance function*

This remark permits a better insight in the meaning of the topographic distance for a grey-tone function. Fig. 2 represents a grey-tone function with a constant slope in each tile of a mosaic image; the mosaic image appears on the support of the function. We will compute the topographic distance along the path (A, B, C, D, E, F, G, H). The projection of this path on the support is the path (a, b, c, d, e, f, g, h). The portion of the path between a and b belongs to a zone of constant slope α. The topographical length of the portion (A, B) will be equal to the geodesic distance of the corresponding projection (a, b) multiplied by the weight α. If the slopes on the successive portions ab, bc, cd, de, ef, fg and gh are respectively α, β, γ, δ, ε, ζ and η, then the topographical distance along the path (AH) is the weighted sum of $\alpha * d(\mathrm{a},\mathrm{b}) + \beta * d(\mathrm{b},\mathrm{c}) + \gamma * d(\mathrm{c},\mathrm{d}) + \delta * d(\mathrm{d},\mathrm{e}) + \varepsilon * d(\mathrm{e},\mathrm{f}) + \zeta * d(\mathrm{f},\mathrm{g}) + \eta * d(\mathrm{g},\mathrm{h})$. From D to E, the topographic distance will be 0, since the slope δ is equal to 0. From F to G the slope is infinite and the distance from f to g is equal to 0; in this case $\delta * d(\mathrm{f},\mathrm{g}) = f(\mathrm{f}) - f(\mathrm{g})$.

3. Topographical distance and catchment basins in the digital space

The digital space always brings the same problems: infinitely small structuring elements do not exist on a digital grid. Furthermore, the neighborhood relations are poor. It is nevertheless possible to reach good approximations of the watershed line.

3.1. *The topographical distance function*

Let us consider a grey-tone function f from $\mathbb{Z}^n$ to $\mathbb{Z}$ with its support $\mathrm{supp}(f)$. Let G be the underlying grid, which can be of any type, square or hexagonal in 2 dimensions, cubic, centred cubic or face centred cubic in 3 dimensions. We associate to G a *neighborhood graph* U. U is a subset of $\mathbb{Z}^n \times \mathbb{Z}^n$ defined by $(a, b) \in U$ iff a and b are neighbors.

We call $N_U(p)$ the set of neighbors of a pixel p, with respect to U and to $\mathrm{supp}(f)$: $N_U(p) = \{p' \in \mathbb{Z}^2, (p, p') \in U \cap \mathrm{supp}(f)\}$. The subset $B(p)$ of all pixels of $N_U(p)$ which are at a distance 1 of p is called elementary ball of size 1. The erosion of the function f by this ball is the elementary erosion εf.

Definition 5. A path π of cardinal n between two pixels p and q on the grid G is an n-tuple of pixels $(p_1, p_2, \ldots, p_n)$, such that $p_1 = p$, $p_n = q$, and $\forall_i \in [1, n-1]$, $(p_i, p_{i+1}) \in G$.

The length of the path π is defined by $l(\pi) = \sum_i \mathrm{dist}(p_i, p_{i+1})$.

The slope between two pixels p and p' for $f(p') < f(p)$ is defined by $\mathrm{slope}(p, p') = (f(p) - f(p'))/\mathrm{dist}(p, p')$.

Definition 6. The set of lower neighbors of p, for which $\mathrm{slope}(p, p')$ is maximal is written as $\Gamma(p)$. The value of this maximal slope is called lower slope of the function at the point p:

$$\mathrm{LS}(p) = \max\left(\frac{f(p) - f(p')}{\mathrm{dist}(p, p')}\right)$$

for $p' \in N_U(p)$ and $f(p') < f(p)$.

Remark. If $N_G(p) = B(p)$ then the lower slope is simply computed with the help of the elementary erosion: $\mathrm{LS}(p) = (I - \varepsilon) f(p) = \partial f(p)$.

Remark. The most frequently used neighborhood graphs are the trivial ones, corresponding to

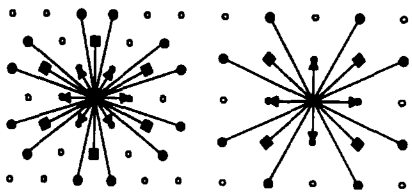

Fig. 3. Basic Chamfer neighborhood relations for the square and the hexagonal grid.

4-connectivity on the square grid or 6-connectivity on the hexagonal grid. It may however be more complex, yielding 8 or even 16 neighbors on the square grid and 12 or 24 neighbors on the hexagonal grid. These neighborhoods are called Chamfer neighborhoods and are illustrated in Fig. 3. They are classically used for the construction of distance functions [4]. The distances from each neighbor to the central point are approximated with integer values. For instance, in square connectivity with 16 neighbors, the distances are 5, 7 and 11. In 6 connectivity, for 12 neighbors the weights they are 4 and 7. For pixels p and p' which are not immediate neighbors, slope(p, p') will only be computed if there exists a descending path between p and p'.

Definition 7. The mapping $\Gamma(p)$ allows us to define an oriented graph V, as a subgraph of the neighborhood graph U:

$$(p, p') \in V \Leftrightarrow p' \in \Gamma(p).$$

Definition 8. We may now define the cost for walking on the topographic surface from one position $f(p_{i-1})$ to a neighbor position $f(p_i)$:

$$f(p_{i-1}) > f(p_i) \rightarrow$$

$$\text{cost}(p_{i-1}, p_i) = \text{LS}(p_{i-1}) * \text{dist}(p_{i-1}, p_i),$$

$$f(p_{i-1}) < f(p_i) \rightarrow$$

$$\text{cost}(p_{i-1}, p_i) = \text{LS}(p_i) * \text{dist}(p_{i-1}, p_i),$$

$$f(p_{i-1}) = f(p_i) \rightarrow$$

$$\text{cost}(p_{i-1}, p_i) = \frac{\text{LS}(p_{i-1}) + \text{LS}(p_i)}{2} * \text{dist}(p_{i-1}, p_i).$$

Definition 9. Let f be a grey-tone function and π a path $(p_1 = p, p_2, \ldots, p_n = q)$ between two pixels p and q inside supp(f). Then the π-topographical distance between p and q on f along the path π is the weighted distance defined by

$$T_f^{\pi}(p, q) = \sum_{i>1} \text{cost}(p_{i-1}, p_i).$$

Definition 10. The topographical distance between two pixels p and q is defined as the minimal π-topographical distance between the two pixels p and q among all pathes π between p and q inside supp(f):

$$T_f(p, q) = \inf(T_f^{\pi}(p, q)).$$

Proposition 3. *The topographical distance is an écart but not a distance, since the separability of distances is not satisfied:*
(a) *positivity:* $T_f(p, q) \geqslant 0$,
(b) *symmetry:* $T_f(p, q) = T_f(q, p)$,
(c) *no separability:* $T_f(p, q) = 0$ *does not imply that* p *and* q *are the same pixel. The topographical distance between two pixels belonging to the interior of the same plateau is equal to zero, since the lower slope of all pixels of the interior of a plateau is equal to* 0. *We will see later how to transform this ecart into a distance by introducing an auxiliary order relation inside the plateaus.*
(d) *triangular inequality:* $T_f(p, q) \leqslant T_f(p, r) + T_f(r, q)$.

Proof. (a) and (b) follow immediately from the definition of the π-topographical distance. The triangular inequality also is satisfied: if π_1 (respectively π_2) is a path for which $T_f(p, r)$ (respectively $T_f(p, q)$) is minimal, then the minimum along all pathes between p and q is smaller than the minimum along the juxtaposition of the pathes π_1 and π_2. □

Remark. We have $p_{i-1} \in \Gamma(p_i)$ if and only if $\text{cost}(p_{i-1}, p_i) = f(p_i) - f(p_{i-1})$ In the other cases the value of cost is higher.

Definition 11. A pixel q belongs to the upstream of a pixel p, if there exists a path π of steepest slope between p and q: $\pi = (p_1 = p, p_2, \ldots, p_n = q)$ and $\forall_i; p_{i-1} \in \Gamma(p_i)$.

The preceding remark implies the following proposition.

Proposition 4. *Let p and q be two points such that $f(p) < f(q)$. Then we have the following equivalence*:

$$T_f(p,q) = f(q) - f(p) \iff q \in \text{upstream}(p).$$

In all other situations we have $T_f(p,q) > |f(p) - f(q)|$ which has as a consequence:

Proposition 5. *The topographic distance between a pixel x and the regional minimum m_i in the depth of its catchment basin is equal to $f(x) - f(m_i)$ and the geodesic line between them is a line of steepest descent.*

3.2. Definition of the catchment basins

The definition of catchment basins and watershed lines remains the same than in the continuous space (Definitions 3 and 4), if one replaces the continuous distance T_D by the discrete distance T_f.

4. Watershed lines and shortest paths algorithms

In Definition 3 we have seen how to assign a cost to each arc of the neighborhood graph. The nodes belonging to the regional minima are assigned an initial cost equal to their altitude. The construction of the catchment basins of a grey-tone function becomes a problem of finding a path with minimal cost between each pixel and a regional minimum. All pixels along a minimal cost path will get the same label as the regional minimum at the origin of the path. This problem of finding a shortest path in a weighted graph is classical in operational research and has been solved 35 years ago. Based on these algorithms, which have been proved to be correct, we will be able to rephrase correct, and for some of them new, algorithms for the construction of the watershed line.

4.1. Flooding from regional minima or flooding from markers

Gradient images possess many regional minima, and only a few are meaningful for the purpose of segmentation. Constructing the watershed line of a gradient image most often yields an oversegmented image. As we have recalled in the introduction, the solution to this problem is the introduction of a set of markers, from which the relief of the gradient image is flooded. The algorithms we will now see, however, describe how to construct the watershed line associated to the set of regional minima. How to introduce markers in that case? There exists a classical solution to this problem, due to Beucher and Meyer, called homotopy modification. It is presented in the next section.

4.1.1. Modification of the homotopy of the gradient image

The classical solution [9, 2] consists in replacing the original gradient g by a new function g'. The images g and g' differ only by their regional minima: the only minima of g' correspond to the set of markers. These minima have a value equal to 0. All minima of g have been filled in. The algorithm for performing this operation is simple and relies on a grey-tone reconstruction. All details can be found in [6, 2]. After this transformation, any classical algorithm for constructing the watershed from a set of regional minima can be applied.

4.2. The algorithm of Moore [10]

4.2.1. The distance from a node to all others

The aim of the algorithm published in 1957 by Moore [10] is to compute the shortest path from a node to all other nodes. The principle is the following. The nodes for which the length of the shortest path is known are ordered according to their length. The node with the lowest value is expanded and the shortest paths of its neighbors computed. The algorithm solves the problem of finding the distance from a node to all other nodes. It will be easy to derive a version finding the catchment basins from this algorithm. We follow here the presentation given in [5].

Let (X, W) be a graph, where X represents the nodes and W the arcs between nodes. l_{ij} is the cost associated with the arc (i, j) for $(i, j) \in W$.
Let $\pi^*(i)$ be the minimal length of all paths between 1 and i; in particular $\pi^*(1) = 0$.

The algorithm will proceed in $N - 1$ iterations. At the beginning of each iteration, the set of nodes is partitioned in two subsets, S and $\bar{S} = X - S$. For the first iteration $1 \in S$.

Each node i of X is given a label $\pi(i)$ satisfying the following property:
- if $i \in S$, $\pi(i) = \pi^*(i)$,
- if $i \in \bar{S}$, $\pi(i) = \min_{k \in S, (k,i) \in W} (\pi(k) + l_{ki})$.

The value $\pi(i)$ for $i \in \bar{S}$ gives the minimal length of the pathes between 1 and i under the condition that all nodes of the path except i are included in the set S.

The correctness of the algorithm is due to the following lemma.

Lemma. *Let j be the node of $\bar{S}$ verifying $\pi(j) = \min_{i \in \bar{S}} \pi(i)$. Then $\pi^*(j) = \pi(j)$.*

Proof. There exists a path from 1 to j with the length $\pi(j)$. In order to show that this path is the shortest possible, let us consider another path μ; we may cut it into two parts: the first part μ_1 starts at 1 and ends at h, where h is the first node of $\bar{S}$ reached, and μ_2 the remaining part of the path. Then by definition of j we have $\pi(h) \geqslant \pi(j)$. And the length of μ_1 verifies length$(\mu_1) \geqslant \pi(h)$. On the other hand we have length$(\mu_2) \geqslant 0$. Putting everything together we get:

$$\text{length}(\mu) = \text{length}(\mu_1) + \text{length}(\mu_2) \geqslant \pi(h) \geqslant \pi(j).$$

The algorithm of Moore is then the following:

(a) Initialization:
$\bar{S} = \{2, 3, \ldots, N\}$,

$$\pi(1) = 0, \qquad \pi(i) = \begin{cases} l_{1i} & \text{if } (1, i) \in W, \\ \infty & \text{if not.} \end{cases}$$

(b) Select the node $j \in \bar{S}$ satisfying $\pi(j) = \min_{i \in \bar{S}} \pi(i)$.
Do: $\bar{S} \leftarrow \bar{S} - \{j\}$. If $|\bar{S}| = 0$ END; else go to (c).

(c) For each i such that $(i, j) \in W$ and $i \in \bar{S}$, do $\pi(i) \leftarrow \min(\pi(i), \pi(j) + l_{ji})$ and return to (b).

4.2.2. Computing the catchment basins by integration

The algorithm of Moore can easily be adapted in order to compute the catchment basins of a greytone function f for which we know the set of regional minima (m_i). We adopt the same notations than in Section 3. G represents all nodes of the grid and U the neighborhood relations. Each regional minimum has a label which is expanded to all pixels belonging to the catchment basin of this minimum.

(a) Initialization. For all pixels of the regional minima the minimal distance is known and is equal to the altitude of the regional minima: $\forall x \in m_i$, $\pi(x) = f(x)$. For all other pixels z, we put $\pi(z) = \infty$. The inside pixels of the regional minima, i.e. the pixels without a higher neighbor are put into the set S; all other pixels including the pixels of the inner boundary ∂m_i of the regional minima ($x \in \partial m_i \Leftrightarrow x \in m_i$ and $\exists z$, $(x, z) \in U$, $f(z) > f(x)$) are put in the set $\bar{S}$.

(b) Select the pixel $x \in \bar{S}$ for which $\pi(x) = \min_{z \in \bar{S}} \pi(z)$. Remove x from $\bar{S}$: $\bar{S} \leftarrow \bar{S} - \{x\}$.
If $\bar{S}$ is empty: END, else go to (c).

(c) For each neighbor z of x inside the set $\bar{S}: z \in \bar{S} \cap N_U(x)$ do:
if $\pi(z) < \pi(x) + \text{cost}(x, z)$ then $\pi(z) = \pi(x) + \text{cost}(x, z)$ and label (z) = label(x).
Return to (b).

4.2.3. Computing the catchment basins by hill climbing

The case where the function f is not known and the integration of the gradient has to be done at the same time that the watershed line is computed is not frequent. The most common case where this happens is when the gradient is equal to 1, yielding a distance function. In the other situations the function f is known. We have seen in Proposition 5 that the geodesics between a regional minimum and the pixels of its catchment basin are lines of steepest descent. It is easy to check on the preceding algorithm that it really is the case. This remark permits to construct a simpler algorithm using the mapping Γ defined in Definition 6.

4.2.3.1. A simplified algorithm. In the general case, the distances are not the same. This corresponds to chamfer neighborhoods. We have seen that in this case, each pixel z gets its value from its neighbor belonging to $\Gamma(z)$: during the expansion of x, all its neighbors belonging to the inverse mapping $\Gamma^{-1}(x)$ without a label get the label of x. The algorithm then becomes the following.

(a) Initialization. For all pixels of the regional minima the label is known. The inside pixels of the regional minima are not to be expanded and belong to the set S. The boundary points of the regional minima are to be expanded and belong to $\bar{S}$ together with all pixels outside the regional minima.
(b) Select the pixel $x \in \bar{S}$ for which $f(x) = \min_{z \in \bar{S}} f(z)$. Remove x from $\bar{S}: \bar{S} \leftarrow \bar{S} - \{x\}$. If $\bar{S}$ is empty: END, else go to (c).
(c) For each pixel z belonging to $\Gamma^{-1}(z) \cap \bar{S}$ without label do: label(z) = label(x). Return to (b).

Remark. If the neighborhood $N_U(p)$ is restricted to the first neighbors $B(p)$ of p, then $\Gamma(p)$ includes only the lowest neighbors of p, i.e. the first neighbors to be expanded in Moores algorithm. This implies that when a pixel z is expanded, the set $\Gamma^{-1}(z)$ includes all neighbors of z which did not get a label earlier.

The step (c) of the preceding algorithm may be simplified:

(c′) For each neighbor z of x belonging to $\bar{S}$ without label do: label(z) = label(x). Return to (b).

4.2.4. The blindness of the watershed algorithms on plateaus

4.2.4.1. Cause of the blindness. Step (b) consists in selecting the pixel with the smallest value $f(x)$. We discuss here the problem where several pixels are candidates for selection, because they have the same value. This situation happens particularly on plateaus of the function f. On such plateaus there are no guidelines for the progression of the flood; this means there is no uniqueness of the watershed line. It is clear that among all possible solutions, some solutions make more sense than others: one would like the watershed line to be situated on some median line of the plateaus rather than on the borders.

4.2.4.2. Hierarchical queues. For the Moore type algorithms there exists a simple and elegant solution: if one adopts an adequate data structure for the storage of all pixels with labels and the retrieval of the pixel with the lowest value, one gets both a high speed of treatment and a correct placement of the watershed lines on the plateaus. One such data structure is called ordered queues and has the following feature [6]: the pixels are ordered according to their altitude; for each class of altitude a file is used as storage medium with the principle ‘first in first out’. The pixels at the boundary of the plateaus enter the corresponding file as their lower neighbors are expanded. These pixels will be the first to come out the file and during their expansion the pixels at a distance 2 of the boundary enter the file. With this simple mechanism, the pixels are treated in the order of increasing distances to the lower border of the plateaus. This feature corrects for the blindness of the topographical distance within plateaus.

4.2.4.3. Arrowing. Another solution leaves more freedom; it consists in modifying the lowest neighbor graph V introduced in Definition 6 within the plateaus, i.e. for all pairs of pixels (p, p') for which $\text{cost}(p, p') = 0$. For such a pair of pixels, an oriented arc will be created from p to p' if the geodesic distance to the lower border of the plateau is greater for p than for p'. This may be done easily by constructing the distance function on each plateau to its lower border. The lower neighbor graph is constructed for this distance function. The union of both graphs will be called completed graph and written CV.

4.3. The algorithm of Berge

4.3.1. The distance from a node to all others

Berge published in 1958 an important shortest path algorithm [1], very different in nature from the algorithm of Moore. It relies on the following remark. Let $\pi(i)$ be the shortest path between 1 and i. Let (i, j) be the arc of length l_{ij} linking the nodes i and j. The shortest path between the node 1 and the node j may pass through the node i; this implies: $\pi(j) \leqslant \pi(i) + l_{ij}$. The function attributing to each node its weighted distance to 1 should verify the relation: $\pi(j) - \pi(i) \leqslant l_{ij}$. The algorithm of Berge is then the following:

(a) Initialization: $\pi(1) = 0$ and $\forall_i \neq 1$, $\pi(i) = 0$.
(b) Find an arc (i, j), for which $\pi(j) - \pi(i) > l_{ij}$. If such an arc does not exist: END
(c) Write $\pi(j) = \pi(i) + l_{ij}$ and return to (b).

Remark 1. This algorithm gives to each node i decreasing values $\pi(i)$. Since the values are lower bounded by 0, the algorithm converges. It is easy to prove that the limit is effectively the length of the shortest path to the node 1.

Remark 2. The great interest of this algorithm is that it does not impose any order on the treatment of the pixels. A good efficiency will be obtained with the classical scanning modes of sequential algorithms: a forward scanning followed by a backward scanning. These scannings are repeated until stability is reached. This makes the hardware design much simpler: it is easy to construct image memories allowing fast forward and backward scannings. It is not so easy to construct memories whose access is driven by a hierarchical queue.

4.3.2. Flooding from the set of minima and solving the eikonal equation

Let (m_i) be the set of regional minima of f. We know for each pixel its lower slope $\mathrm{LS}(x)$; we have seen in Section 3.1 how to derive from it the $\mathrm{cost}(x, y)$ associated to each arc (x, y). Each regional minimum possesses a label $\mathrm{lab}(m_i)$.

We already noticed that topographical distance is completely blind on the plateaus. For the algorithm of Moore this was not so much of a problem, since the hierarchical queues impose a natural order within the plateaus. However, in the case of the algorithm of Berge, the position of the watershed line does depend upon the scanning order. For instance, for a forward scanning order the position of the watershed line will always be shifted to the bottom rightside of the image. This is not acceptable for functions which may have large plateaus. A mechanism has then to be found to correct this bias. We will introduce an additional control function identical after convergence of the algorithm with the distance function to the lower border of each plateau. Let θ be this function.

The inside points of plateaus have a lower slope equal to 0.

(a) Initialization. For each pixel x belonging to a regional minimum: $\pi(x) = 0$. For all other points $\pi(x) = \infty$. For each pixel x for which $\mathrm{LS}(x) = 0$, do $\theta(x) = \infty$. For all other pixels: $\theta(x) = 0$.
(b) Repeat a forward raster scanning followed by an inverse raster scanning, and apply to the center pixel x the following treatment:
{
For each neighbor y of x, belonging to the future of x for the current scanning do:
if $\pi(y) \geqslant \pi(x) + \mathrm{cost}(x, y)$ then $\pi(y) = \pi(x) + \mathrm{cost}(x, y)$ and $\mathrm{lab}(y) = \mathrm{lab}(x)$
if $\{\mathrm{cost}(x, y) = 0$ and $\theta(y) \geqslant \theta(x) + \mathrm{dist}(x, y)\}$ then $\{\theta(y) = \theta(x) + \mathrm{dist}(x, y)$ and $\mathrm{lab}(y) = \mathrm{lab}(x)\}$
} until the output of a complete cycle of forward and backward raster scanning is identical to its input.

Remark. If the lower slope LS is strictly positive everywhere outside the regional minima, then the function f has no plateaus and in the preceding algorithm, any line containing ϑ may be skipped.

4.3.3. Flooding from the set of minima when the function itself is known

We introduce the same oriented subgraph V as in Section 4.1.3. This orientation will guide the propagation of the labels of the regional minima; the propagation within the plateaus being guided by the distance function on the plateaus underlying the arrowing of V.

(a) Initialization. For all pixels x outside regional minimum: $\mathrm{lab}(x) = \infty$.
(b) Repeat a forward raster scanning followed by an inverse raster scanning, and apply to the center pixel the following treatment:
{
For each neighbor y of x, belonging to the future of for the current scanning do:
if $\mathrm{lab}(y) \neq \mathrm{lab}(x)$ and $y \in \Gamma(x)$ then $\mathrm{lab}(y) = \mathrm{lab}(x)$.

} until no further modification occurs during a complete cycle of forward and backward raster scanning.

4.4. Superiority of the chamfer watershed lines

4.4.1. Precision of the watershed line

If the neighborhood $N_G(p)$ counts more pixels than the first neighbors, we get a higher precision in the placement of the watershed line. The classical neighborhoods are shown in Fig. 3. The integral of a constant function equal to 1 with such a neighborhood yields the classical Chamfer distance functions. Applied to grey-tone image, the same algorithm will permit a more precise placement of the watershed line and yield a 'chamfer watershed line'.

4.4.2. Comparison between the different algorithms

Integrating a constant function with a constant grey-tone yields a distance function to all sets serving as limit conditions. Using larger neighborhoods yield chamfer distance functions which are known to be superior to the distance functions based solely on the first neighbors.

Applied on classical gradient images, the superiority of the chamfer algorithm is not as visible than on distance functions. This is due to the fact, that most real images, on which the watershed is computed share the same feature: the watershed line is perpendicular to the lines of steepest slope. When flooding such a relief, the floods come from two opposite directions. Under such circumstances, first neighbor algorithms make a reasonable job. For this reason, Matheron suggested a much tougher test model, where the angle between the watershed line and the lines of steepest descent take all possible values, from perpendicular to almost parallel. In the case where quasi-parallelism occurs, most of the algorithms fail, except the chamfer algorithms.

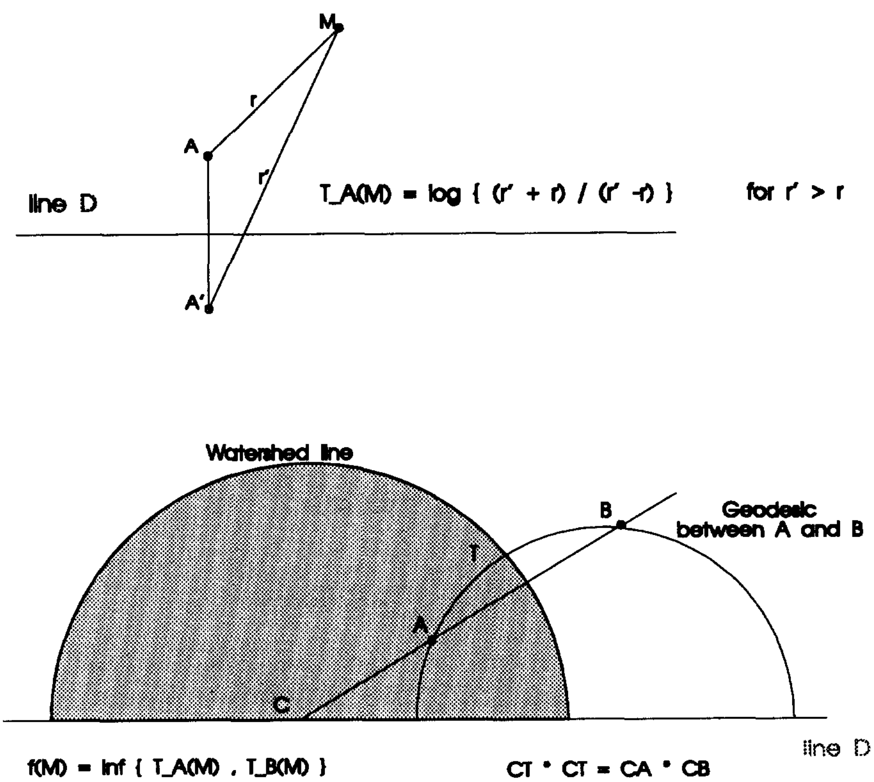

Fig. 4. (a) Potential associated to a point A, and its value at a point M. (b) The grey-tone function f is the minimum of the potentials associated to points A and B. In grey the catchment basin of the point A.

The test model is constructed as follows. One defines first a potential function associated to a point A. This potential is defined in a half-plane, limited by a straight line D. One considers the point A in the half-plane and its symmetric A' by the straight line D (see Fig. 4). The potential at the point M associated to the point M is then defined by

$$T\_A(M) = \log\left\{\frac{(r'+r)}{(r'-r)}\right\},$$

where $r = |AM|$ and $r' = |A'M|$.

This potential defines a kind of distance to the point M. This distance has a uniformly decreasing gradient, as the distance to the line D decreases. If y is the distance to the line D, the modulus of the gradient of f is $|\mathrm{grad}(f)| = 1/y$. Let us consider a path between two pixels x and y, belonging to the support of the function. As the path moves away from the line D, the topographical distance along the path will decrease, whereas the geodesic distance along the path remains the same.

If we consider now two points A and B, we may define the potentials T_A and T_B associated to these points. We will construct the watershed line of

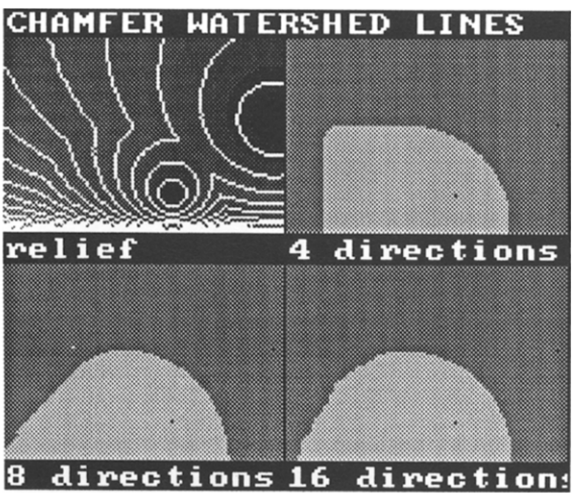

Fig. 5. Construction of the function f of the definition. (a) The level lines of the function f are indicated in white. (b) Construction of the watershed line with a neighborhood of 4 pixels. (c) Neighborhood of 8 pixels. (d) Neighborhood of 16 pixels.

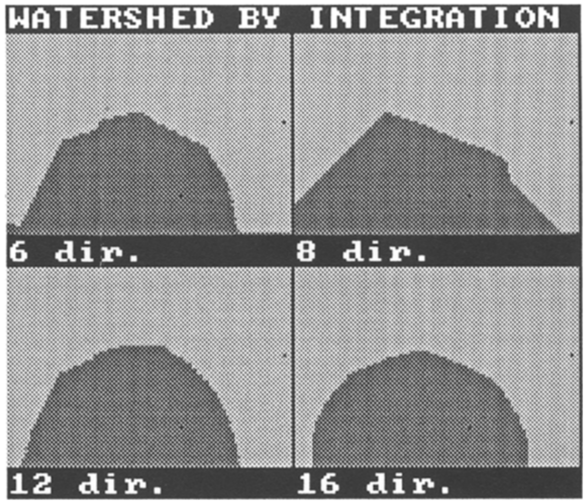

Fig. 6. Construction of the catchment basins by integrating the modulus of the function $|\mathrm{grad}(f)| = |1/y|$. (a) and (c) are in hexagonal raster, with 6 and 12 neighbors, respectively. (b) and (d) are in square raster, with 8 and 16 neighbors, respectively.

the function $f(M) = \inf\{T\_A(M), T\_B(M)\}$ (see Fig. 4(b). The points A and B are the two only regional minima of the function f. The position of the watershed line of f can be computed theoretically: it is a half circle centered on the line D, at the intersection point C the lines D and AB. The radius $\rho = |CT|$ of the circle verifies: $CT{*}CT = CA{*}CB$.

Since the modulus of the gradient of the function f is known, two ways for testing the algorithms are possible: (a) construct directly the function f and detect the catchment basin, (b) construct the catchment basins by integrating the modulus of the gradient, which is known. The superiority of the larger neighborhoods is blatant as shown in Figs. 5 and 6.

References

[1] C. Berge, *Théorie des graphes et ses applications*, Dunod, Paris, 1958.

[2] S. Beucher, Segmentation d'images et morphologie mathématique, Thèse Ecole des Mines, 1990.

[3] S. Beucher and C. Lantuéjoul, "Use of watersheds in contour detection", *Proc. Workshop on Image Processing, CCETT/IRISA*, Rennes, France, September 1979, pp. 2.1–2.12.

[4] G. Borgefors, "Distance transformations in digital images", *CVGIP*, Vol. 14, 1980, pp. 227–248.
[5] M. Gondran and M. Minoux, *Graphes et algorithmes*, Eyrolles, Paris, 1979, chapitre 4: Les algèbres de chemins, pp. 65–101.
[6] F. Meyer, "Un algorithme optimal de ligne de partage des eaux", *Actes du 8ème Congres AFCET*, Lyon-Villeurbanne, France, 1991, pp. 847–859.
[7] F. Meyer, "Integrals and gradients of images", *Proc. SPIE, Image Algebra and Morphological Image Processing III*, San Diego, CA, Vol. 1769, 1992, pp. 200–211.
[8] F. Meyer, "Integrals, gradients and watershed lines", *Proc. Mathematical Morphology and its applications to Signal Processing*, Barcelona, May 1993, pp. 70–75.
[9] F. Meyer and S. Beucher, "Morphological segmentation", *JVCIP*, Vol. 1, No. 1, September 1990, pp. 21–46.
[10] E.F. Moore, "The shortest path through a maze", *Proc. Internat. Symp. on Theory of Switching*, 1957, Annals of the computation laboratory of Harvard University 30, 1959, pp. 285–292.
[11] L. Najman and M. Schmitt, "Definition and some properties of the watershed of a continuous function", *Proc. Mathematical Morphology and its Applications to Signal Processing*, Barcelona, May 1993, pp. 75–81.
[12] F. Preteux, "On a distance function approach for grey-level mathematical morphology", in: E.R. Dougherty, ed., *Mathematical Morphology in Image Processing*, Marcel Dekker, New York, 1993, pp. 323–351.
[13] P.W. Verbeek and J.H. Verwer, "Shading from shape, the eikonal equation solved by grey-weighted distance transformation", *Pattern Recognition Lett.*, Vol. 11, 1990, pp. 681–690.