



HAL
open science

Ultrametric watersheds

Laurent Najman

► **To cite this version:**

Laurent Najman. Ultrametric watersheds. 9th International Symposium on Mathematical Morphology (ISMM'09), Aug 2009, Groningen, Netherlands. pp.181-192. hal-00622405

HAL Id: hal-00622405

<https://hal.science/hal-00622405v1>

Submitted on 12 Sep 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ultrametric watersheds

Laurent Najman

Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, Equipe A3SI,
ESIEE Paris, France

Abstract. We study hierarchical segmentation in the framework of edge-weighted graphs. We define ultrametric watersheds as topological watersheds null on the minima. We prove that there exists a bijection between the set of ultrametric watersheds and the set of hierarchical edge-segmentations.

Introduction

This paper¹ is a contribution to a theory of hierarchical image segmentation in the framework of edge-weighted graphs. Image segmentation is a process of decomposing an image into regions which are homogeneous according to some criteria. Intuitively, a hierarchical segmentation represents an image at different resolution levels.

In this paper, we introduce a subclass of edge-weighted graphs that we call ultrametric watersheds. Theorem 9 states that there exists a one-to-one correspondance, also called a bijection, between the set of indexed hierarchical edge-segmentations and the set of ultrametric watersheds. In other words, to any hierarchical edge-segmentation (whatever the way the hierarchy is built), it is possible to associate a representation of that hierarchy by an ultrametric watershed. Conversely, from any ultrametric watershed, one can infer a indexed hierarchical edge-segmentation.

Following [1], we can say that, independently of its theoretical interest, such a bijection theorem is useful in practice. Any hierarchical segmentation problem is a priori heterogeneous: assign to an edge-weighted graph an indexed hierarchy. Theorem 9 allows such classification problem to become homogeneous: assign to an edge-weighted graph a particular edge-weighted graph called ultrametric watershed. Thus, Theorem 9 gives a meaning to questions like: which hierarchy is the closest to a given edge-weighted graph with respect to a given measure or distance?

The paper is organised as follow. Related works are examined in section 1. We introduce segmentation on edges in section 2, and in section 3, we adapt the topological watershed framework from the framework of graphs with discrete weights on the nodes to the one of graphs with real-valued weights on the edges. We then define (section 4) hierarchies and ultrametric distances. The last part of the paper (section 5) introduces hierarchical edge-segmentations and ultrametric

¹ This work was partially supported by ANR grant SURF-NT05-2.45825

watersheds, the main result being the existence of a bijection between these two sets (theorem 9).

Apart from Theorems 2 and 3, and to the best of the author's knowledge, all the properties and theorems formally stated in this paper are new. Proofs of the various properties and theorems will be given in an extended version [2] to be published in a journal paper.

1 Related works

1.1 Hierarchical clustering

From its beginning in image processing, hierarchical segmentation is thought of as a particular instance of hierarchical classification [3]. One of the fundamental theorems for hierarchical clustering states that there exists a one-to-one correspondence between the set of indexed hierarchical classification and a particular subset of dissimilarity measures called ultrametric distances; This theorem is generally attributed to Johnson [4], Jardine *et al.* [5] and Benzécri [3]. Since then, numerous generalisations of that bijection theorem have been proposed (see [1] for a recent review).

Our main theorem is an extension to hierarchical edge-segmentation of this fundamental hierarchical clustering theorem.

1.2 Hierarchical segmentation

There exist many methods for building a hierarchical segmentation [6], which can be divided in three classes: bottom-up, top-down, and split-and-merge. A recent review of some of those approaches can be found in [7]. A useful representation of hierarchical segmentations was introduced in [8] under the name of *saliency map*. This representation has been used (under several names) by several authors, for example for visualisation purposes [9] or for comparing hierarchies [10].

In this paper, we show that any saliency map is an ultrametric watershed, and conversely.

1.3 Watersheds

For bottom-up approaches, a generic way to build a hierarchical segmentation is to start from an initial segmentation and progressively merge regions together [11]. Often, this initial segmentation is obtained through a watershed [8, 12, 13]. See [14] for a recent review of these notions in the context of mathematical morphology.

Among many others [15], topological watershed [16] is an original approach to watershed that modifies a map (e.g., a grayscale image) while preserving the connectivity of each lower cross-section (see fig. 2). It has been proved [16, 17] that this approach is the only one that preserves altitudes of the passes (named connection values in this paper) between regions of the segmentation.

Pass altitudes are fundamental for hierarchical schemes [8]. On the other hand, topological watersheds may be thick. A study of the properties of different kinds of graphs with respect to the thinness of watersheds can be found in [18, 19]. An interesting framework is that of edge-weighted graphs, where watersheds are naturally thin; furthermore, in that framework, a subclass of topological watersheds satisfies both the drop of water principle and a property of global optimality [20].

In this paper, we translate topological watersheds from the framework of node-weighted-graphs to the one of edge-weighted graphs, and we identify ultrametric watersheds, a subclass of topological watersheds that is interesting for hierarchical edge-segmentation.

2 Segmentation on edges

This paper is settled in the framework of edge-weighted graphs. Following the notations of [21], we present some basic definitions to handle such kind of graphs.

We define a *graph* as a pair $X = (V, E)$ where V is a finite set and E is composed of unordered pairs of V , *i.e.*, E is a subset of $\{\{x, y\} \subseteq V \mid x \neq y\}$. We denote by $|V|$ the cardinal of V , *i.e.*, the number of elements of V . Each element of V is called a *vertex or a point (of X)*, and each element of E is called an *edge (of X)*. If $V \neq \emptyset$, we say that X is *non-empty*.

As several graphs are considered in this paper, whenever this is necessary, we denote by $V(X)$ and by $E(X)$ the vertex and edge set of a graph X .

A graph X is said *complete* if $E = V(X) \times V(X)$.

Let X be a graph. If $u = \{x, y\}$ is an edge of X , we say that x and y are *adjacent (for X)*. Let $\pi = \langle x_0, \dots, x_\ell \rangle$ be an ordered sequence of vertices of X , π is a *path from x_0 to x_ℓ in X (or in V)* if for any $i \in [1, \ell]$, x_i is adjacent to x_{i-1} . In this case, we say that x_0 and x_ℓ are *linked for X* . We say that X is *connected* if any two vertices of X are linked for X .

Let X and Y be two graphs. If $V(Y) \subseteq V(X)$ and $E(Y) \subseteq E(X)$, we say that Y is a *subgraph of X* and we write $Y \subseteq X$. We say that Y is a *connected component of X* , or simply a *component of X* , if Y is a connected subgraph of X which is maximal for this property, *i.e.*, for any connected graph Z , $Y \subseteq Z \subseteq X$ implies $Z = Y$.

Let X be a graph, and let $S \subseteq E(X)$. The *graph induced by S* is the graph whose edge set is S and whose vertex set is made of all points which belong to an edge in S , *i.e.*, $(\{x \in V(X) \mid \exists u \in S, x \in u\}, S)$.

Important remark. *Throughout this paper $G = (V, E)$ denotes a connected graph, and the letter V (resp. E) will always refer to the vertex set (resp. the edge set) of G . We will also assume that $E \neq \emptyset$.*

Let $S \subseteq E$. In the following, when no confusion may occur, the graph induced by S is also denoted by S .

Typically, in applications to image segmentation, V is the set of picture elements (pixels) and E is any of the usual adjacency relations, *e.g.*, the 4- or 8-adjacency in 2D [22].

If $S \subset E$, we denote by \overline{S} the complementary set of S in E , i.e., $\overline{S} = E \setminus S$.

A set $C \subset E$ is an (edge-)cut (of G) if each edge of C is adjacent to two different nonempty connected components of \overline{C} .

A graph S is called an (edge-)segmentation (of G) if $\overline{E(S)}$ is a cut.

Any connected component of a segmentation S is called a region (of S).

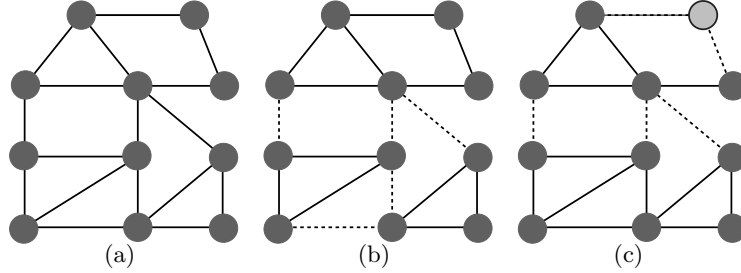


Fig. 1. Illustration of edge-segmentation and edge-cut. (a) A graph X . (b) An edge-segmentation of X ; the set of dotted-lines edges is the associated edge-cut of X . (c) A subgraph of X which is not an edge-segmentation of X : the grey point is isolated.

The previous definitions of cut and segmentation (illustrated on fig. 1) are not the usual ones. In particular, Prop. 1.(i) below states that there is no isolated point in an edge-segmentation. If we need an isolated point x , it is always possible to replace x with an edge $\{x', y'\}$. Furthermore, isolated points are often noise in an image.

It is interesting to state the definition of a segmentation from the point of view of vertices of the graph. A graph X is said to be *spanning* (for V) if $V(X) = V$. We denote by ϕ the map that associates, to any $X \subset G$, the graph $\phi(X) = \{V(X), \{\{x, y\} \in E \mid x \in V(X), y \in V(X)\}\}$. We observe that $\phi(X)$ is maximal among all subgraphs of G that are spanning for $V(X)$, it is thus a closing on the lattice of subgraphs of G [23]. We call ϕ the edge-closing.

Property 1 A graph $S \subseteq G = (V, E)$ is an edge-segmentation of G if and only if

- (i) The graph induced by $E(S)$ is S ;
- (ii) S is spanning for V ;
- (iii) for any connected component X of S , $X = \phi(X)$.

3 Topological watershed

3.1 Edge-weighted graphs

We denote by \mathcal{F} the set of all mappings from E to \mathbb{R}^+ and we say that any mapping in \mathcal{F} *weights* the edges of G . For any $F \in \mathcal{F}$, the pair (G, F) is called

an *edge-weighted graph*. Whenever no confusion can occur, we will denote the edge-weighted graph (G, F) by F .

For applications to image segmentation, we will assume that the altitude of u , an edge between two pixels x and y , represents the dissimilarity between x and y (e.g., $F(u)$ equals the absolute difference of intensity between x and y ; see [24] for a more complete discussion on different ways to set the mapping F for image segmentation). Thus, we suppose that the salient contours are located on the highest edges of (G, F) .

Let $\lambda \in \mathbb{R}^+$ and $F \in \mathcal{F}$, we define $F[\lambda] = \{v \in E \mid F(v) \leq \lambda\}$. The graph (induced by) $F[\lambda]$ is called a *(cross)-section* of F . A connected component of a section $F[\lambda]$ is called a *(level λ) component* of F .

We define $\mathcal{C}(F)$ as the set composed of all the pairs $[\lambda, C]$, where $\lambda \in \mathbb{R}^+$ and C is a component of the graph $F[\lambda]$. We call *altitude of $[\lambda, C]$* the number λ . We note that one can reconstruct F from $\mathcal{C}(F)$; more precisely, we have:

$$F(v) = \min\{\lambda \mid [\lambda, C] \in \mathcal{C}(F), v \in E(C)\}$$

For any component C of F , we set $h(C) = \min\{\lambda \mid [\lambda, C] \in \mathcal{C}(F)\}$. We define $\mathcal{C}^*(F)$ as the set composed by all $[h(C), C]$ where C is a component of F . The set $\mathcal{C}^*(F)$, called the *component tree* of F [25, 26], is a finite subset of $\mathcal{C}(F)$ that is widely used in practice for image filtering.

A *(regional) minimum* of F is a component X of the graph $F[\lambda]$ such that for all $\lambda_1 < \lambda$, $F[\lambda_1] \cap E(X) = \emptyset$. We remark that a minimum of F is a subgraph of G and not a subset of the points of G ; we also remark that any minimum X of F is such that $|V(X)| > 1$.

We denote by $\mathcal{M}(F)$ the graph whose vertex set and edge set are, respectively, the union of the vertex sets and edge sets of all minima of F .

3.2 Topological watersheds on edge-weighted graphs

Let X be a subgraph of G . An edge $u \in \overline{E(X)}$ is said to be *W-simple* (for X) (see [16]) if X has the same number of connected components as $X + u = (V(X) \cup u, E(X) \cup \{u\})$. An edge u such that $F(u) = \lambda$ is said to be *W-destructible* (for F) with lowest value λ_0 if there exists λ_0 such that, for all $\lambda_1, \lambda_0 < \lambda_1 \leq \lambda$, u is W-simple for $F[\lambda_1]$ and if u is not W-simple for $F[\lambda_0]$.

A *topological watershed* (on G) is a mapping that contains no W-destructible edges.

A mapping F' is a *topological thinning* (of F) if:

- $F' = F$, or if
- there exists a mapping F'' which is a topological thinning of F and there exists an edge u W-destructible for F'' with lowest value λ such that $\forall v \neq u, F'(v) = F''(v)$ and $F'(v) = \lambda_0$, with $\lambda \leq \lambda_0 < F''(v)$.

An illustration of a topological watershed can be found in fig. 2.

The *connection value* between $x \in V$ and $y \in V$ is the number

$$F(x, y) = \min\{\lambda \mid x \in V(C), y \in V(C), [\lambda, C] \in \mathcal{C}(F)\} \quad (1)$$

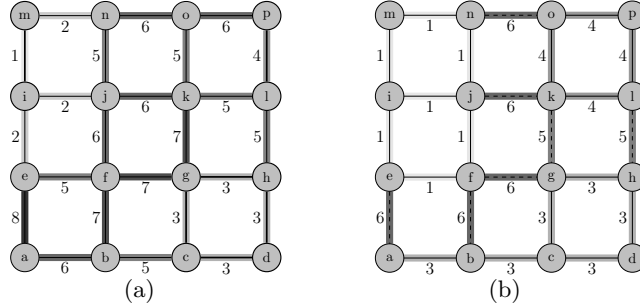


Fig. 2. Illustration of topological watershed. (a) An edge-weighted graph F . (b) A topological watershed of F . The minima of (a) are $(\{m, i\}, \{p, l\}, \{g, h\}, \{c, d\}, \{g, c\}, \{h, d\})$ and are in bold in (a).

In other words, $F(x, y)$ is the altitude of the lowest element $[\lambda, C]$ of $\mathcal{C}(F)$ such that x and y belong to C (rule of the least common ancestor).

Two points x and y are *separated (for F)* if $F(x, y) > \max\{\lambda_1, \lambda_2\}$, where λ_1 (resp. λ_2) is the altitude of the lowest element $[\lambda_1, c_1]$ (resp. $[\lambda_2, c_2]$) of $\mathcal{C}(F)$ such that $x \in c_1$ (resp. $y \in c_2$). The points x and y are λ -*separated (for F)* if they are separated and $\lambda = F(x, y)$.

The mapping F' is a *separation of F* if, whenever two points are λ -separated for F , they are λ -separated for F' .

If X and Y are two subgraphs of G , we set $F(X, Y) = \min\{F(x, y) \mid x \in X, y \in Y\}$.

Theorem 2 (Restriction to minima [16]). *Let $F' \leq F$ be two elements of \mathcal{F} . The mapping F' is a separation of F if and only if, for all distinct minima X and Y of $\mathcal{M}(F)$, we have $F'(X, Y) = F(X, Y)$.*

A graph X is *flat (for F)* if for all $u, v \in E(X)$, $F(u) = F(v)$. If X is flat, the *altitude of X* is the number $F(X)$ such that $F(X) = F(v)$ for any $v \in E(X)$.

We say that F' is a *strong separation of F* if F' is a separation of F and if, for each $X' \in \mathcal{M}(F')$, there exists $X \in \mathcal{M}(F)$ such that $X \subseteq X'$ and $F(X) = F(X')$.

Theorem 3 (strong separation [16]). *Let F and F' in \mathcal{F} with $F' \leq F$. Then F' is a topological thinning of F if and only if F' is a strong separation of F .*

In other words, topological thinnings are the only way to obtain a watershed that preserves connection values.

In the framework of edge-weighted graphs, topological watersheds allows for a simple characterization.

Theorem 4. *A mapping F is a topological watershed if and only if:*

- (i) $\mathcal{M}(F)$ is a segmentation of G ;

(ii) for any edge $v = \{x, y\}$, if there exist X and Y in $\mathcal{M}(F)$, $X \neq Y$, such that $x \in V(X)$ and $y \in V(Y)$, then $F(v) = F(X, Y)$.

Note that if F is a topological watershed, then for any edge $v = \{x, y\}$ such that there exists $X \in \mathcal{M}(F)$ with $x \in V(X)$ and $y \in V(X)$, we have $F(v) = F(X)$.

4 Hierarchies and ultrametric distances

Let Ω be a finite set. A *hierarchy* H on Ω is a set of parts of Ω such that

- (i) $\Omega \in H$
- (ii) for every $\omega \in \Omega$, $\{\omega\} \in H$
- (iii) for each pair $(h, h') \in H^2$, $h \cap h' \neq \emptyset \implies h \subset h'$ or $h' \subset h$.

An *indexed hierarchy* on Ω is a pair (H, μ) , where H denotes a given hierarchy on Ω and μ is a positive function, defined on H and satisfying the following conditions:

- (i) $\mu(h) = 0$ if and only if h is reduced to a singleton of Ω ;
- (ii) if $h \subset h'$, then $\mu(h) < \mu(h')$.

A distance d , in general, obeys the triangular inequality $d(\omega_1, \omega_2) \leq d(\omega_1, \omega_3) + d(\omega_3, \omega_2)$ where ω_1, ω_2 and ω_3 are any three points of the space. An *ultrametric distance* (on Ω) is a function d from $\Omega \times \Omega$ to \mathbb{R}^+ such that $d(\omega_1, \omega_2) = 0$ if and only if $\omega_1 = \omega_2$, such that $d(\omega_1, \omega_2) = d(\omega_2, \omega_1)$ and such that d obeys ultrametric inequality $d(\omega_1, \omega_2) \leq \max(d(\omega_1, \omega_3), d(\omega_2, \omega_3))$ for all $\omega_1, \omega_2, \omega_3$. The ultrametric inequality [27] is stronger than the triangular inequality.

A *partition* of Ω is a collection (Ω_i) of non-empty subsets of Ω such that any element of Ω is exactly in one of these subsets. Note that any given partition of the set Ω induces a large number of trivial ultrametric distances: $d(\omega_1, \omega_1) = 0$, $d(\omega_1, \omega_2) = 1$ if $\omega_1 \in \Omega_i, \omega_2 \in \Omega_j, i \neq j$, and $d(\omega_1, \omega_2) = a$ if $i = j, 0 < a < 1$. The general connection between indexed hierarchies and ultrametric distances was proved by Benzécri [3] and Johnson [4]. This result states that there is a one-to-one correspondance between indexed hierarchies and ultrametric distances both defined on the same set. Indeed, associated with each indexed hierarchy (H, μ) on Ω is the following ultrametric distance:

$$d(\omega_1, \omega_2) = \min\{\mu(h) \mid \omega_1 \in h, \omega_2 \in h, h \in H\}. \quad (2)$$

In other words, the distance $d(\omega_1, \omega_2)$ between two elements ω_1 and ω_2 in Ω is given by the smallest element in H which contains both ω_1 and ω_2 . Conversely, each ultrametric distance d is associated with one and only one indexed hierarchy.

Observe the similarity between eq. 2 and eq. 1. Indeed, connection value is an ultrametric distance on V whenever $F > 0$. More precisely, we have the following property.

Property 5 *Let $F \in \mathcal{F}$. Then $F(X, Y)$ is an ultrametric distance on $\mathcal{M}(F)$. If furthermore, $F > 0$, then $F(x, y)$ is an ultrametric distance on V .*

Let Ψ be the mapping on \mathcal{F} such that for any $F \in \mathcal{F}$ the map $\Psi(F)$ and for any edge $\{x, y\} \in E$, $\Psi(F)(\{x, y\}) = F(x, y)$. It is straightforward to see that $\Psi(F) \leq F$, that $\Psi(\Psi(F)) = \Psi(F)$ and that if $F' \leq F$, $\Psi(F') \leq \Psi(F)$. Thus Ψ is an opening on the lattice (\mathcal{F}, \leq) [28]. We remark that the subset of strictly positive maps that are defined on the complete graph $(V, V \times V)$ and that are open with respect to Ψ is the set of ultrametric distances on V . The mapping Ψ is known under several names, in particular the one of subdominant ultrametric and the one of ultrametric opening. It is well known that Ψ is associated to the simplest method for hierarchical classification called single linkage clustering [5, 29], closely related to Kruskal's algorithm [30] for computing a minimum spanning tree.

Thanks to Th. 4, we observe that if F is a topological watershed, then $\Psi(F) = F$. However, an ultrametric distance d may have plateaus, and thus the weighted complete graph $(V, V \times V, d)$ is not always a topological watershed. Nevertheless, those results underline that topological watersheds are related to hierarchical classification, but not yet to hierarchical edge-segmentation; the study of such relations is the subject of the rest of the paper.

5 Hierarchical edge-segmentations, saliency and ultrametric watersheds

Informally, a hierarchical segmentation is a hierarchy made of connected regions. However, in our framework, a segmentation is not a partition, and as the union of two disjoint connected subgraphs of G is not a connected subgraph of G , the formal definition is slightly more involved. A *hierarchical (edge-)segmentation (on G)* is an indexed hierarchy (H, μ) on the set of regions of a segmentation S of G , such that for any $h \in H$, $\phi(\cup_{X \in h} X)$ is connected (ϕ being the edge-closing defined in section 2).

For any $\lambda \geq 0$, we denote by $H[\lambda]$ the graph induced by $\{\phi(\cup_{X \in h} X) | h \in H, \mu(h) \leq \lambda\}$. The following property is an easy consequence of the definition of a hierarchical segmentation.

Property 6 *Let (H, μ) be a hierarchical segmentation. Then for any $\lambda \geq 0$, the graph $H[\lambda]$ is a segmentation of G .*

Property 5 implies that the connection value defines a hierarchy on the set of minima of F . If F is a topological watershed, then by Th. 4, $\mathcal{M}(F)$ is a segmentation of G , and thus from any topological watershed, one can infer a hierarchical segmentation. However, $F[\lambda]$ is not always a segmentation: if there exists a minimum X of F such that $F(X) = \lambda_0 > 0$, for any $\lambda_1 < \lambda_0$, $F[\lambda_1]$ contains at least two connected components X_1 and X_2 such that $|V(X_1)| = |V(X_2)| = 1$. Note that the value of F on the minima of F is not related to the position of the divide nor to the associated hierarchy of minima/segmentations. This leads us to introduce the following definition.

A map $F \in \mathcal{F}$ is an *ultrametric watershed* if F is a topological watershed, and if furthermore, for any $X \in \mathcal{M}(F)$, $F(X) = 0$.

Property 7 A map F is an ultrametric watershed if and only if for all $\lambda \geq 0$, $F[\lambda]$ is a segmentation of G .

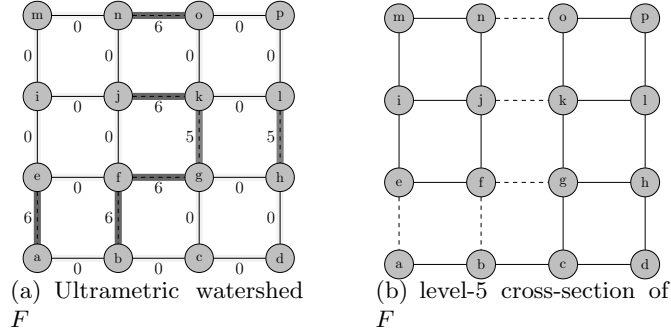


Fig. 3. An example of an ultrametric watershed F and a cross-section of F .

This property is illustrated in fig. 3.

By definition of a hierarchy, two elements of H are either disjoint or nested. If furthermore (H, μ) is a hierarchical segmentation, the graphs $\overline{E(H[\lambda])}$ can be stacked to form a map. We call *saliency map* [8] the result of such a stacking, *i.e.* a saliency map is a map F such that there exists (H, μ) a hierarchical segmentation with $F(v) = \min\{\lambda | v \in E(H[\lambda])\}$.

Property 8 A map F is a saliency map if and only if F is an ultrametric watershed.

A corollary of property 8 states the equivalence between hierarchical segmentations and ultrametric watersheds. The following theorem is the main result of this paper.

Theorem 9. There exists a bijection between the set of hierarchical edge-segmentations on G and the set of ultrametric watersheds on G .

As there exists a one-to-one correspondance between the set of indexed hierarchies and the set of ultrametric distances, it is interesting to search if there exists a similar property for the set of hierarchical segmentations. Let d be the ultrametric distance associated to a hierarchical segmentation (H, μ) . We call *ultrametric contour map (associated to (H, μ))* the map d_E such that:

1. for any edge $v \in E(H[0])$, then $d_E(v) = 0$;
2. for any edge $v = \{x, y\} \in \overline{E(H[0])}$, $d_E(v) = d(X, Y)$ where X (resp. Y) is the connected component of $H[0]$ that contains x (resp. y).

Property 10 A map F is an ultrametric watershed if and only if F is the ultrametric contour map associated to a hierarchical segmentation.

6 Conclusion

Fig. 4 is an illustration of the application of the framework developed in this paper to a classical hierarchical segmentation scheme based on attribute opening [8, 14, 25]. Fig. 5 shows some of the differences between applying such scheme and applying a classical morphological segmentation scheme, *e.g.* attribute opening followed by a watershed [12]. As watershed algorithms generally place watershed lines in the middle of plateaus, the two schemes give quite different results.

It is important to note that most of the algorithms proposed in the literature to compute saliency maps are not correct, often because they rely on wrong connection values or because they rely on thick watersheds where merging regions is difficult. Future papers will propose novel algorithms (based on the topological watershed algorithm [31]) to compute ultrametric watersheds, with proof of correctness.

On a more theoretical level, this work can be pursued in several directions.

- We will study lattices of watersheds [32] and will bring to that framework recent approaches like scale-sets [9] and other metric approaches to segmentation [10]. For example, scale-sets theory considers a rather general formulation of the partitioning problem which involves minimizing a two-term-based energy, of the form $\lambda C + D$, where D is a goodness-of-fit term and C is a regularization term, and proposes an algorithm to compute the hierarchical segmentation we obtain by varying the λ parameter. We can hope that the topological watershed algorithm [31] can be used on a specific energy function to directly obtain the hierarchy.
- Subdominant theory (mentioned at the end of section 4) links hierarchical classification and optimisation. In particular, the subdominant ultrametric d' of a dissimilarity d is the solution to the following optimisation problem for $p < \infty$:

$$\min\{\|d - d'\|_p^p \mid d' \text{ is an ultrametric distance and } d' \leq d\}$$

It is certainly of interest to search if topological watersheds can be solutions of similar optimisation problems.

- Several generalisations of hierarchical clustering have been proposed in the literature [1]. An interesting direction of research is to see how to extend in the same way the topological watershed approach, for example for allowing regions to overlap.

References

1. Barthélemy, J.P., Brucker, F., Osswald, C.: Combinatorial optimization and hierarchical classifications. *4OR: A Quarterly Journal of Operations Research* **2**(3) (2004) 179–219
2. Najman, L.: Ultrametric watersheds: a bijection theorem for hierarchical edge-segmentation. Technical Report IGM 2009-07, Université Paris-Est, Institut Gaspard Monge (2009)

3. Benzécri, J.: *L'Analyse des données: la Taxinomie*. Volume 1. Dunod (1973)
4. Johnson, S.: Hierarchical clustering schemes. *Psychometrika* **32** (1967) 241–254.
5. Jardine, N., Sibson, R.: *Mathematical taxonomy*. Wiley (1971)
6. Pavlidis, T.: Hierarchies in structural pattern recognition. *Proceedings of the IEEE* **67**(5) (May 1979) 737–744
7. Soille, P.: Constrained connectivity for hierarchical image decomposition and simplification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **30**(7) (July 2008) 1132–1145
8. Najman, L., Schmitt, M.: Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(12) (December 1996) 1163–1173
9. Guignes, L., Cocquerez, J.P., Men, H.L.: Scale-sets image analysis. *International Journal of Computer Vision* **68**(3) (2006) 289–317
10. Arbeláez, P.A., Cohen, L.D.: A metric approach to vector-valued image segmentation. *International Journal of Computer Vision* **69**(1) (2006) 119–126
11. Pavlidis, T. In: *Structural Pattern Recognition*. Volume 1 of Springer Series in Electrophysics. Springer (1977) 90–123 segmentation techniques, chapter 4–5.
12. Meyer, F., Beucher, S.: Morphological segmentation. *Journal of Visual Communication and Image Representation* **1**(1) (September 1990) 21–46
13. Meyer, F.: Morphological segmentation revisited. In: *Space, Structure and Randomness*. Springer (2005) 315–347
14. Meyer, F., Najman, L.: Segmentation, arbre de poids minimum et hiérarchies. In Najman, L., Talbot, H., eds.: *Morphologie mathématique 1: approches déterministes*. Lavoisier, Paris (2008) 201–233
15. Roerdink, J.B.T.M., Meijster, A.: The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae* **41**(1-2) (2001) 187–228
16. Bertrand, G.: On topological watersheds. *J. Math. Imaging Vis.* **22**(2-3) (May 2005) 217–230
17. Najman, L., Couprie, M., Bertrand, G.: Watersheds, mosaics and the emergence paradigm. *Discrete Appl. Math.* **147**(2-3) (2005) 301–324
18. Cousty, J., Bertrand, G., Couprie, M., Najman, L.: Fusion graphs: merging properties and watersheds. *J. Math. Imaging Vis.* **30**(1) (January 2008) 87–104
19. Cousty, J., Najman, L., Bertrand, G., Couprie, M.: Weighted fusion graphs: merging properties and watersheds. *Discrete Appl. Math.* **156**(15) (August 2008) 3011–3027
20. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: minimum spanning forests, and the drop of water principle. *IEEE Trans. Pattern Anal. Mach. Intell.* (2009) To appear.
21. Diestel, R.: *Graph Theory*. Graduate Texts in Mathematics. Springer (1997)
22. Kong, T., Rosenfeld, A.: Digital topology: Introduction and survey. *Comput. Vision Graph. Image Process.* **48**(3) (1989) 357–393
23. Cousty, J., Najman, L., Serra, J.: Some morphological operators in graph spaces. In: *ISMM 09*. (2009) To appear
24. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: thinnings, shortest-path forests and topological watersheds. *IEEE Trans. Pattern Anal. Mach. Intell.* (2009) to appear.
25. Salembier, P., Oliveras, A., Garrido, L.: Anti-extensive connected operators for image and sequence processing. *IEEE Trans. on Image Proc.* **7**(4) (April 1998) 555–570

26. Najman, L., Couprie, M.: Building the component tree in quasi-linear time. *IEEE Transactions on Image Processing* **15**(11) (2006) 3531–3539
27. Krasner, M.: Espaces ultramétrique. *C.R. Acad. Sci. Paris* **219** (1944) 433–435
28. Leclerc, B.: Description combinatoire des ultramétries. *Mathématique et sciences humaines* **73** (1981) 5–37
29. Gower, J., Ross, G.: Minimum spanning tree and single linkage cluster analysis. *Appl. Stats.* **18** (1969) 54–64
30. Kruskal, J.B.: On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society* **7** (February 1956) 48–50
31. Couprie, M., Najman, L., Bertrand, G.: Quasi-linear algorithms for the topological watershed. *J. Math. Imaging Vis.* **22**(2-3) (2005) 231–249
32. Cousty, J., Najman, L., Serra, J.: Raising in watershed lattices. In: 15th IEEE ICIP'08, San Diego, USA (October 2008) 2196–2199

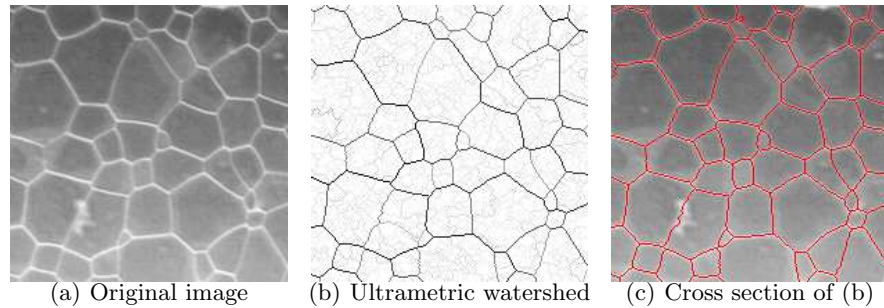


Fig. 4. Example of ultrametric watershed.

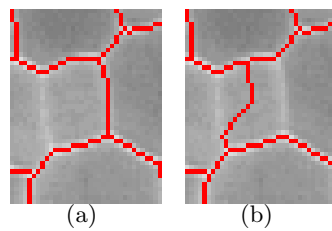


Fig. 5. Zoom on a comparison between two watersheds of a filtered version of the image 4.a. Morphological filtering tends to create large plateaus, and both watersheds (a) and (b) are possible, but only (a) is a subset of a watershed of 4.a. No hierarchical scheme will ever give a result as (b).