



Energy Partitions and Image Segmentation

PABLO ANDRÉS ARBELÁEZ AND LAURENT D. COHEN

CEREMADE, UMR CNRS 7534 Université Paris Dauphine, Place du maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France

arbelaez@ceremade.dauphine.fr

cohen@ceremade.dauphine.fr

Abstract. We address the issue of low-level segmentation for real-valued images. The proposed approach relies on the formulation of the problem in terms of an energy partition of the image domain. In this framework, an energy is defined by measuring a pseudo-metric distance to a source point. Thus, the choice of an energy and a set of sources determines a tessellation of the domain. Each energy acts on the image at a different level of analysis; through the study of two types of energies, two stages of the segmentation process are addressed. The first energy considered, the path variation, belongs to the class of energies determined by minimal paths. Its application as a pre-segmentation method is proposed. In the second part, where the energy is induced by an ultrametric, the construction of hierarchical representations of the image is discussed.

Keywords: segmentation, distance transform, path variation, ultrametrics, extrema mosaic, mathematical morphology, image enhancement

1. Introduction

Image segmentation is a fundamental issue in the field of computer vision. Its great complexity may be explained by the fact that structuring visual information into *meaningful regions* involves a layer of semantic understanding of the image content. As pointed out by a recent study [28], even the evaluation of human segmentation is a challenging problem, where subjectivity and recognition often prevail over the raw physical data. Therefore, the integration of prior knowledge appears as an essential step in the creation of any high-level computer vision system.

However, a first task is the extraction of the image structure provided by the interaction between low-level cues. The present paper addresses the segmentation problem at this pre-cognitive stage of perception, focusing on the case of monochrome images. For this purpose, we propose an extension of the well known Voronoi tessellation [3, 35] to pseudo-metric spaces, which we call an *energy partition* of the image domain. In this framework, an energy is defined by measuring

the shortest pseudo-metric distance to a fixed source point. Then, given a set of sources, a tessellation of the domain is constructed by the assignation of every point to the closest source. Hence, in this context, the segmentation issue is transferred to the definition of an appropriated energy, based on the image data, and the selection of a set of sources.

The interpretation of vision as an information processing task [27] suggests the use of a multilevel strategy to address computer vision problems. Following this idea, two types of energies are studied. Each energy acts on the image at a specific level of analysis, allowing to treat a particular stage of the segmentation process.

The first type of energies considered is issued from the study of minimal paths. Their use is often appropriated for a local level of analysis of the image. We introduce a particular energy of this class, the *path variation*, which is an extension of the one dimensional total variation for real-valued functions of two variables. Then, choosing the intensity extrema as sources, a simplified version of the original image is obtained. This natural

reconstruction of the image, referred as the *extrema mosaic*, provides a balance between content conservation and simplification.

In the second part, we study a class of energies that present interesting properties for a global level of analysis. These energies are associated to the ultrametrics and their definition is equivalent to the construction of a family of nested partitions of the image domain. Starting from the extrema mosaic, an energy that extracts the global contrast of the original image is defined. Then, other ultrametric energies are constructed by using the internal information of the regions as a complement to the contrast. Their application to the creation of hierarchical representations of the image is discussed.

The rest of this paper is organized as follows. The general framework and the notations are introduced in Section 2. Section 3 presents the energies derived from the study of minimal paths. Section 4 is devoted to the study of the path variation and Section 5 presents the extrema mosaic. Section 6 begins by recalling the basic properties of ultrametric spaces; then, a global measure of contrast for the image is constructed. In Section 7, the approach is applied to hierarchical segmentation. Finally, Section 8 contains some concluding remarks.

2. Energy Partitions

In this section, the general framework of energy partitions is presented; basic definitions are recalled and the notations introduced.

Let $\Omega \subset \mathbb{R}^2$ be a compact connected domain in the plane. An application $\delta : \Omega \times \Omega \rightarrow \mathbb{R}^+$ is a **pseudo-metric** for Ω [19] if, for any $x, y, z \in \Omega$, the following conditions are satisfied:

Identity:

$$\delta(x, x) = 0 \quad (1a)$$

Symmetry:

$$\delta(x, y) = \delta(y, x) \quad (1b)$$

Triangle Inequality:

$$\delta(x, y) \leq \delta(x, z) + \delta(z, y) \quad (1c)$$

Note that the only difference with the definition of a metric is that the usual *Separation* axiom was replaced by the weaker condition (1a). Therefore, considering the **equivalence classes** $\hat{x}(\delta) = \{y \in \Omega \mid \delta(x, y) = 0\}$

leads to work directly on the **quotient space** $\hat{\Omega}(\delta) = \{\hat{x}(\delta) \mid x \in \Omega\}$.

Thus, the projection of δ on $\hat{\Omega}(\delta)$, is, by definition, a metric for the quotient space. Note that, if δ is already a metric, then $\hat{\Omega}(\delta)$ coincides with the domain. Additionally, the existence of geodesics for δ is assumed. In the sequel, the value of $\delta(x, y)$ will be referred as the *distance* between the points x and y .

The **energy** induced by a pseudo-metric δ , with respect to a source point $s \in \Omega$, is defined as the single variable application $\delta_s : \Omega \rightarrow \mathbb{R}^+$, that measures the distance to s :

$$\delta_s(x) = \delta(s, x), \quad \forall x \in \Omega.$$

The energy with respect to a set of sources $S = \{s_i\}_{i \in J}$ is given by the minimal individual energy:

$$\delta_S(x) = \inf_{s_i \in S} \delta_{s_i}(x), \quad \forall x \in \Omega.$$

In the presence of multiple sources, a valuable information is provided by the **influence zone** of a source s_i , or the set of points that are closer to s_i than to any other source, in the sense of δ :

$$Z_i = \{x \in \Omega \mid \delta_{s_i}(x) < \delta_{s_j}(x), \quad \forall s_j \in S, j \neq i\}.$$

Thus, the influence zone, or the *zone* for short, is a connected subset of the domain. Their union is noted by:

$$Z(\delta, S) = \bigcup_{i \in J} Z_i.$$

The **medial set** is defined as the complementary set of $Z(\delta, S)$:

$$M(\delta, S) = \{x \in \Omega \mid \exists s_i, s_j \in S : \delta_S(x) = \delta_{s_i}(x) = \delta_{s_j}(x)\}.$$

Hence, a pseudo-metric δ and a set of sources S determine a tessellation of the domain. In the sequel, such a tessellation will be referred as an **energy partition** of Ω and will be denoted by $\Pi(\delta, S)$:

$$\Pi(\delta, S) = \{Z_i\}_{i \in J} \cup \{M(\delta, S)\}.$$

It should be noted that every element of an energy partition is a union of elements of the quotient space $\hat{\Omega}(\delta)$.

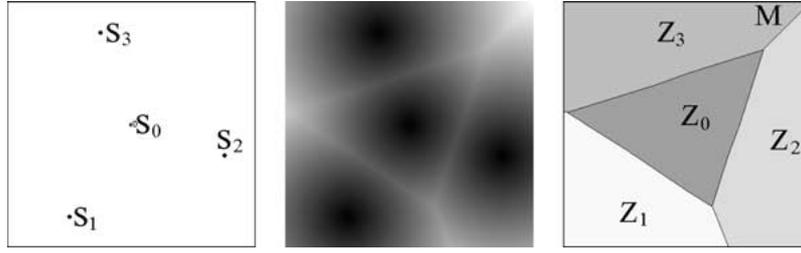


Figure 1. Sources, Euclidean energy and energy partition.

Figure 1 illustrates these definitions with the canonical Euclidean space and a set of four source points, $S = \{s_0, s_1, s_2, s_3\}$, shown on the left. The energy induced by the Euclidean distance, noted by E_S , is displayed on the middle and the energy partition $\Pi(E, S)$ is on the right. Since E is a metric on Ω , the quotient space $\hat{\Omega}(E)$ coincides with the domain. Additionally, the medial set $M(E, S)$ (shown in black) corresponds to the well know Voronoi diagram. Finally, the influence zones $Z(E, S)$ are in this case convex sets.

Two main differences of our approach with the standard framework of Voronoi tessellations [3, 35] should nevertheless be noted. First, by considering pseudo-metrics, a larger class of spaces can be accessed. Second, since we are interested in the application of these notions to image analysis, we discuss the definition of pseudo-metrics that depend on the image.

Therefore, in this context, the image segmentation problem is transferred to the selection of a set of sources and the definition of a relevant energy, or, equivalently, a pseudo-metric, determined by the image data. However, it should be noted that, in practice, digital images are subsampled on the discrete grid. Consequently, important parts of the medial set may fall in the intergrid space. For region based segmentation purposes, an alternative to surround this problem is to consider an energy partition composed only by zones. Thus, the elements of the medial set that would fall in the grid are assigned to one of their neighboring influence zones.

Then, an approximation of the image can be constructed by the assignation of a *model* to represent each influence zone. The model is determined by the distribution of the image values on the zone; simple models are the mean or median value on the influence zone and the source's level. When the model is constant, the valuation of each zone by its model produces a piecewise constant approximation of the image, referred in the sequel as a *mosaic image* (which will be illustrated later in Fig. 4).

3. Path Metrics

A first approach for the definition of a pseudo-metric on an image domain is the study of paths between couples of points.

A **path** between two points $x, y \in \Omega$ is a continuous function $\gamma : [a, b] \rightarrow \Omega$ such that $\gamma(a) = x$ and $\gamma(b) = y$.

The image of γ is then a curve in Ω . Moreover, if $\gamma \in C^1([0, L])$ and an arc-length parameterization of γ is considered (i.e., $\|\dot{\gamma}(l)\| = 1, \forall l \in [0, L]$), then L represents the Euclidean length of the path. The set of paths from x to y is noted by Γ_{xy} and Γ_Ω denotes the set of paths in Ω .

A **length structure** on Ω [15] designates a map $e : \Gamma_\Omega \rightarrow \mathbb{R}^+$ such that:

1. If γ is constant, then $e(\gamma) = 0$.
2. If γ is the juxtaposition of γ_1 and γ_2 , then $e(\gamma) = e(\gamma_1) + e(\gamma_2)$.
3. e is invariant under changes of parameterization.

A length structure e can be used to define a pseudo-metric, which we call the *path metric* associated to e , by considering the minimal value of e along all the paths joining two points:

$$\delta(x, y) = \inf_{\gamma \in \Gamma_{xy}} e(\gamma), \quad \forall x, y \in \Omega.$$

A case of particular interest occurs when e can be expressed as the integral of a potential function $P: \Omega \times \mathcal{S}^1 \rightarrow \mathbb{R}^+$:

$$e(\gamma) = \int_0^L P(\gamma(l), \dot{\gamma}(l)) dl, \quad (2)$$

where l is the arc-length parameter and \mathcal{S}^1 is the unit circle. The energy induced by a metric of this type will be referred as the *potential energy* associated to P . In

this case, $\hat{x}(\delta)$, the class of a point x , corresponds to the largest connected set with null potential that contains x . Thus, if the potential is strictly positive, the quotient space $\hat{\Omega}(\delta)$ coincides with the domain Ω .

Furthermore, if P depends only on the position $\gamma(\cdot)$, the field of geometrical optics provides the following physical interpretation of the energy: the potential $P: \Omega \rightarrow \mathbb{R}^+$ can be seen as a refractive field of indices of an optical medium and δ_s , called the *eikonal* in this context, supplies the optical length of the light rays. Then, the relation between the energy and the potential can be expressed by the *Eikonal Equation*:

$$\|\nabla \delta_s(x)\| = P(x), \quad \forall x \in \Omega, \quad (3)$$

with boundary condition $\delta_s(s) = 0$.

In this particular case, the corresponding metric is known as a **weighted distance transform** and its computation can be performed using Sethian's *Fast Marching* method [8, 39]. Noticing that the information is propagating outwards from the sources, the Fast Marching uses an up-wind scheme to construct a correct approximation of the viscosity solution of (3).

As a simple example, if $P \equiv 1$, then the corresponding weighted distance,

$$\delta(x, y) = \inf_{\gamma \in \Gamma_{xy}} \int_0^L dl,$$

becomes the geodesic distance, or the Euclidean length of the shortest path between x and y . Moreover, if the domain is convex, then δ coincides with the Euclidean distance.

Weighted distances are widely used in computer vision, where the issue becomes the definition of a relevant potential from the image data, in order to address a particular problem. Examples of applications include shape from shading [20], continuous scale morphology [21], shape recovery [25], active contour models [8], differential morphology [26], tubular shape extraction [11] and perceptual grouping [10].

4. The Path Variation

In this section, the path metric obtained by considering the total variation on the paths as the length structure is studied.

4.1. Continuous Domain

For functions of one real variable, the variation is a functional with known properties [16, 34]. It was introduced by Jordan [18] as follows:

Let $f: [a, b] \rightarrow \mathbb{R}$ be a function, $\sigma = \{t_0, \dots, t_n\}$ a finite partition of $[a, b]$ such that $a = t_0 < t_1 < \dots < t_n = b$ and Φ the set of such partitions.

The **variation**, or **total variation**, of f is defined as the (possible infinite) number given by the formula:

$$v(f) = \sup_{\sigma \in \Phi} \sum_{i=1}^n |f(t_i) - f(t_{i-1})|.$$

For real valued functions of two variables, we propose to consider the path metric associated to the total variation, i.e., the minimal variation of the function on all the paths between two points:

Definition 1. The **path variation** of a function $u: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is defined as:

$$V(u)(x, y) = \inf_{\gamma \in \Gamma_{xy}} v(u \circ \gamma), \quad \forall x, y \in \Omega.$$

The space of functions of **bounded path variation** of Ω , noted by $BPV(\Omega)$, is defined as

$$BPV(\Omega) = \{u: \Omega \rightarrow \mathbb{R} \mid \forall x, y \in \Omega, \exists \hat{\gamma} \in \Gamma_{xy} : V(u)(x, y) = v(u \circ \hat{\gamma}) < \infty\}.$$

In the sequel, u is supposed to have bounded path variation. Note that, if $u \in BPV(\Omega)$, then the path variation between any couple of points is not only required to be finite but also to be realized by a path. Hence, Definition 1 assumes the existence of geodesics for V . This assumption is nonetheless reasonable in the case of digital images.

A path γ is said to be *monotone* for u if the function $u \circ \gamma$ is monotone. By definition, if a path is monotone for u , then it is a geodesic for $V(u)$. Conversely, every geodesic for $V(u)$ is a concatenation of monotone paths. Thus, in general, the geodesics of the path variation are not unique.

In the regular framework, the path variation can be characterized as a potential energy, in the sense of (2):

Proposition 1. *If $u \in C^1(\Omega) \cap BPV(\Omega)$, then $V_s(u)$, path variation of u with respect to a source $s \in \Omega$, is the potential energy associated to $P = |D_\tau u|$, the*

absolute value of the directional derivative of u in the tangent direction of the path τ .

Proof: The proof follows from rewriting the definition in the regular case:

If $f \in C^1([a, b])$, then its total variation can be expressed in terms of the derivative [16] by the formula:

$$v(f) = \int_a^b |f'(t)| dt.$$

Thus, if u is a continuously differentiable function, Definition 1 can be reformulated as:

$$V_s(u)(x) = \inf_{\gamma \in \Gamma_{sx}} \int_a^b \left\langle Du(\gamma(t)), \frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|} \right\rangle \|\dot{\gamma}(t)\| dt,$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in \mathbb{R}^2 .

Hence, considering the arc-length parameter l leads to the following expression for the path variation:

$$V_s(u)(x) = \inf_{\gamma \in \Gamma_{sx}} \int_0^L |D_\tau u(\gamma(l))| dl,$$

where τ denotes the unitary tangent vector to the path. \square

The intuitive interpretation of the path variation is illustrated in Fig. 2. Consider a particle moving along the graph of the function u , depicted on the left, and starting at the source s . Then, as shown on the right, the value of $V_s(u)$ at a point x represents the minimal sum of ascents and descents to be travelled to reach x .

Note that, in contrast to the usual extension of the total variation for functions of multiple variables [36, 37], the path variation is defined pointwise. Additionally, $V(u)$ expresses the same geometric notion as the *linear variation*, introduced in [23], though in a different

formulation, as a part of a geometric theory for functions of two variables.

The **component** of u containing x , noted by K_x , designates the connected component of the isolevel set of u that contains x . The level of a component K is noted by $u(K)$ and T_u denotes the space of components of u . The components of a continuous function are closed and pairwise disjoint subsets of Ω . For continuously differentiable functions, the components of the non-singular levels (i.e.: levels t such that $0 \notin \nabla u(u^{-1}(t))$) coincide with the level lines of u and can be represented as closed Jordan curves.

The importance of the components for the path variation is expressed by the following proposition, whose proof is an immediate consequence of Definition 1.

Proposition 2. *If $u \in BPV(\Omega)$, then the quotient space for the path variation $\hat{\Omega}(V(u))$ corresponds to T_u , the space of components of u .*

Thus, $\hat{x}(V(u))$, the equivalence class of a point x for the path variation, coincides with its component K_x . Consequently, each component of $V_s(u)$ is a union of components of u .

Additionally, for a set of sources S , each element of the energy partition $\Pi(V(u), S)$ is a union of components of the function. Hence, the operator that associates $\Pi(V(u), S)$ to the function u is connected [38] and its application simplifies an image u while preserving the contour information. Furthermore, the energy partitions induced by $V(u)$ are invariant under linear contrast changes and rotations. Therefore, the path variation presents a particular interest for image analysis.

Figure 3 illustrates the application of the path variation on two different test functions. On the top row, a smooth function, given by the simple formula $u(x) = c\|x - s_0\|$, where the sources, $S = \{s_0, s_1\}$, correspond to the upper left and the lower right corners

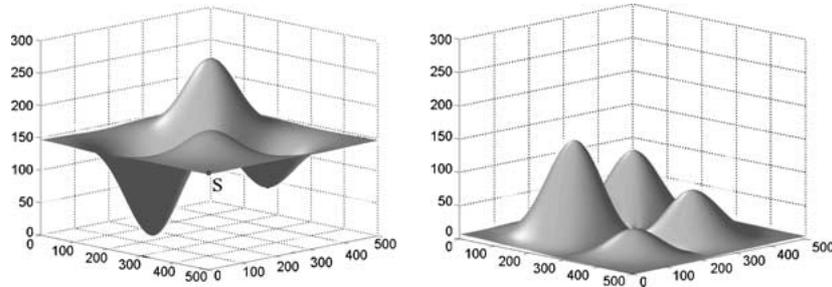


Figure 2. Simple example: graphs of u and $V_s(u)$.

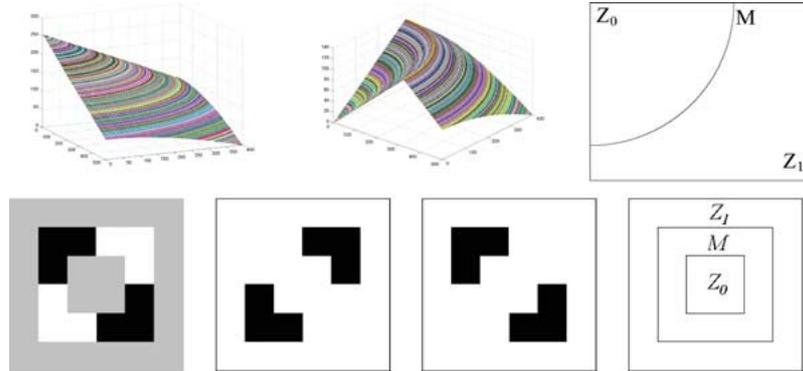


Figure 3. Energy partitions of the path variation (see text).

of a rectangular domain. From left to right, the graph of u , the graph of $V_S(u)$ and the energy partition $\Pi(V(u), S) = Z_0 \cup Z_1 \cup M$ can be observed. In this regular case, the components of the function are nested and coincide with its level lines. The medial set M , shown in black, is the component whose level is the average of the sources' levels.

In contrast, the example on the bottom row of Fig. 3 is the piecewise constant function shown on the left. The corresponding gray levels were set to 0 for the black, 254 for the white and 127 for the gray. The two images on the middle show, in black, the level sets $[u \geq 200]$ and $[u \leq 100]$ respectively. Finally, bottom-right displays the energy partition obtained by taking the two gray components as sources. Notice that the component spaces of the two functions have different topologies: in the first case, T_u is homeomorphic to a segment while, in the second, it is a finite graph. As a consequence, in the second example, although the boundaries of the zones are composed by pieces of level lines, none of the squares determined by the energy partition is a level line of the function.

4.2. Discrete Domain

In this paragraph, a discrete interpretation for the path variation is proposed. For this purpose, we suppose that the image u is sampled on a uniform grid. A first remark is that, since the potential of the path variation in Proposition 1 depends not only on the position but also on the path direction, the Fast Marching method cannot be used for its implementation.

Nevertheless, in a discrete domain, the choice of a digital connectivity determines a notion of component and of vicinity. Thus, T_u , the component space of the

function u , can be represented by an adjacency graph G , where the nodes correspond to discrete components and each link joins two neighboring components. Since the quotient space for $V(u)$ coincides with T_u , we propose to construct the discrete path variation directly on G .

A path on G joining the components of two points p and q is a set of components $\bar{\gamma} = \{K_0, \dots, K_n\}$ such that $p \in K_0, q \in K_n, K_i$ and K_{i-1} are neighbors, $\forall i = 1, \dots, n$. The set of such paths is noted by Γ_{pq}^G . Each element of Γ_{pq}^G corresponds then to a class of discrete paths between p and q .

Thus, the expression of the discrete path variation of u with respect to the source p , evaluated at a point q , becomes:

$$V_p(u)(q) = \min_{\bar{\gamma} \in \Gamma_{pq}^G} \sum_{i=1}^n |u(K_i) - u(K_{i-1})|.$$

Hence, the calculation of $V_p(u)$ is reduced to finding a path of minimal cost on a graph. This classical problem can be solved using a greedy algorithm [12, 24]. The complexity of this straightforward implementation for the path variation is then $O(N \log(N))$, where N is the total number of discrete components of the image. Furthermore, in the case of gray level images, since u takes integer values, the sorting step in the update of the narrow band can be avoided and the complexity is reduced to $O(N)$.

5. The Extrema Mosaic

In this section, the path variation is interpreted as a local measure of contrast and its associated energy partitions are applied to the simplification of monochrome images.

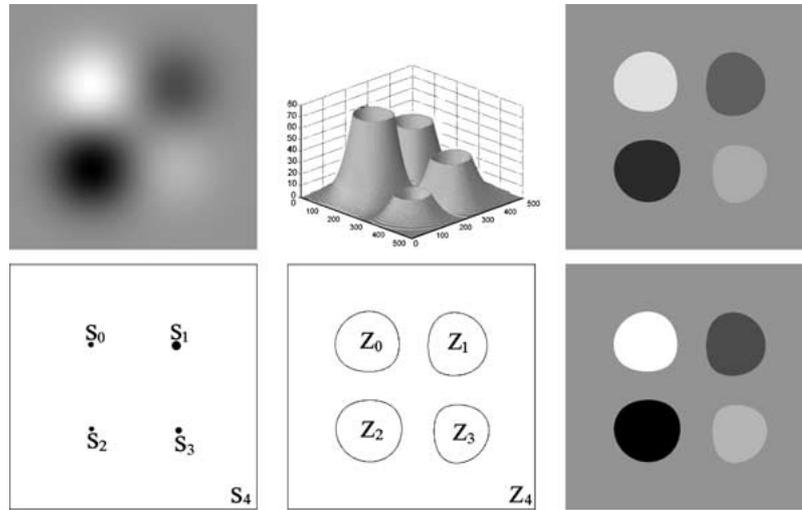


Figure 4. Left: original image and sources. Middle: graph of the energy and energy partition. Right: mosaic images (see text).

Potential energies are often useful for a local level of analysis of the image. Indeed, since their definition (2) is based on an integration along the paths, the partitions determined by this type of pseudo-metrics are very sensitive to the location of the sources.

Therefore, in order to construct a partition based on the path variation, the set of sources should be physically representative of the image content. Figure 4 exemplifies this issue on the smooth image u on top-left. An acceptable segmentation of this “scene” should be composed by four approximately circular regions on a gray background. A solution is to consider a set of sources S composed by the extrema of the four peaks, representing the “features”, and the border of the domain as the source for the background, as shown on bottom-left. The central column displays, on top, the graph of the energy $V_S(u)$ and, on bottom, the associated energy partition $\Pi(V(u), S)$. Finally, the right column shows two mosaic images obtained by taking the median value as the model (top) and the source’s level (bottom). Thus, in the regular framework, the image extrema appear as natural candidates for the sources.

The *extrema partition* of an image u is defined as the energy partition $\Pi(V(u), ext(u))$, where $ext(u)$ denotes the set of extremal components of the intensity. A mosaic image determined by the extrema partition will be called an *extrema mosaic* of u .

Note that, if two extrema can be joined by a monotone path, then the path variation between them coincides with the absolute value of their gray levels’ dif-

ference. Thus, the path variation can be interpreted as a local measure of contrast in the image. Additionally, the application of the extrema mosaic produces a concentration of the contrast information in the boundaries of the zones.

Figures 5 and 6 illustrate the application of our approach on a real image. Figure 5 shows the original image of size 512×512 and the extrema mosaic, with the value of the image at the extremum as the zone model. Figure 6 displays, on top, the detail of size 50×50 marked on Fig. 5 and, on bottom, the graph of the detail.

The simplification performed by the method is expressed by the reduction in the number of components; in this case, the original image has 83430 components (in 8-connectivity) and the mosaic image has only 1253. This example illustrates four properties of the extrema mosaic. First, the previously discussed enhancement of the contrast information can be seen on the detail of the butterfly’s wings. Second, a reduction of the blur in the background, caused by the absorption of blurred contours and transition components by neighboring zones. Third, note how the energy partition preserves the contour information and, particularly, important characteristics of edges such as corners and junctions. Finally, it can be observed that, in spite of the simplification, the resolution of the original image is conserved in the mosaic.

In summary, the choice of the path variation as the energy and the spatial distribution of the intensity extrema determine an energy partition where a

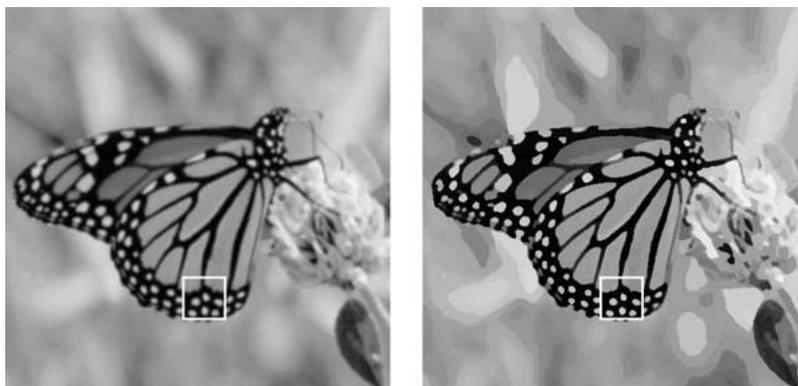


Figure 5. Original image and extrema mosaic.

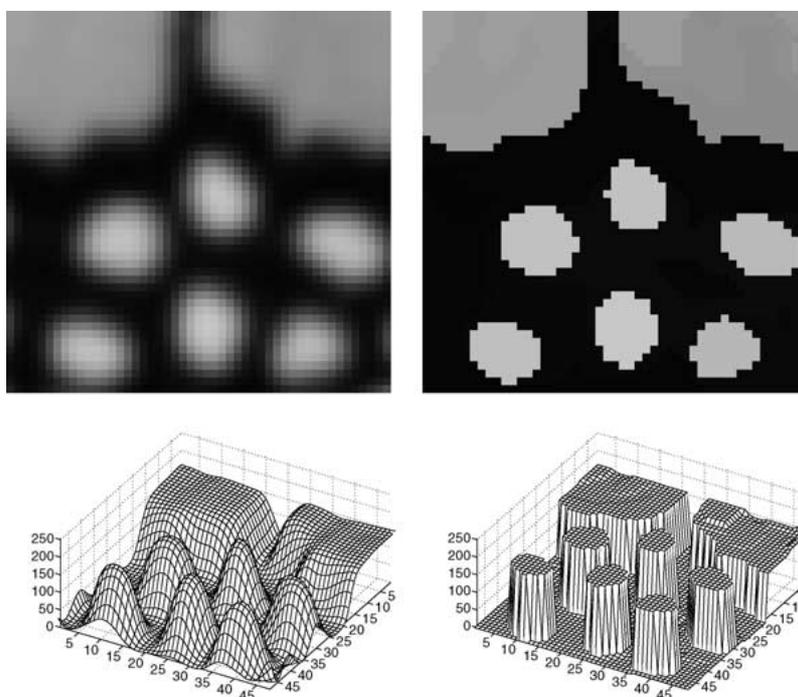


Figure 6. Detail of Fig. 5. Left: Original image. Right: extrema mosaic (see text).

compromise between content conservation and simplification is obtained. Therefore, the extrema partition can be interpreted as a decomposition in elemental zones or as a first level of abstraction of the image information. In [2], we discussed the application of this approach, combined with a nonlinear diffusion filtering, to segmentation problems. In the next sections, this natural reconstruction of the image is used as a parameter-free pre-segmentation method, where the contrast information is enhanced.

6. Ultrametrics

As pointed out in the previous section, the path variation is an interesting tool for the extraction of the local geometric structure of an image. However, when an energy partition with a small number of zones is required, its sensitivity to the location of the sources may become a drawback.

Indeed, for a coarse level of analysis of the image, an energy partition $\Pi(\delta, S)$ should be invariant if a source

$s_i \in S$ is replaced by another point $s'_i \in Z_i$. Unfortunately, this property is generally not satisfied by potential energies. In the next sections, a type of energies that fulfill this invariance requirement is studied.

6.1. Definitions

The energies considered in this part are associated to a class of metrics called the ultrametrics. These energies present a particular interest for segmentation purposes, as they are closely related to the families of nested partitions of the image domain. In this paragraph, some basic properties of ultrametric geometry are recalled.

A **ultrametric** is a metric for which the usual triangle inequality (1c) is replaced by the stronger relation:

$$\delta(x, y) \leq \max\{\delta(x, z), \delta(z, y)\}, \quad \forall x, y, z \in \Omega. \quad (1c')$$

From a geometric point of view, the previous *ultrametric inequality* can be interpreted as follows: all the triangles in an ultrametric space are either isosceles or equilateral.

Furthermore, as a consequence of (1c'), the structure of neighborhoods in an ultrametric space differs significantly from the usual Euclidean space. Firstly, all the points in a ball of radius r and center x , $B_r(x)$, can be considered as the center:

$$\forall y \in B_r(x), \quad B_r(y) = B_r(x). \quad (4)$$

Additionally, two non-disjoint ultrametric balls are always concentric:

$$B_r(x) \cap B_{r'}(y) \neq \emptyset \Rightarrow B_r(x) \subseteq B_{r'}(y) \quad \text{or} \\ B_{r'}(y) \subseteq B_r(x). \quad (5)$$

Therefore, the set of all the balls of a fixed radius r determines a partition of the domain.

Furthermore, (4) and (5) imply that considering the sets of ultrametric balls of radius r , when r is increasing, produces a family of nested partitions of the domain. Additionally, the radii of the balls determine a **stratification index** for the family of partitions and the resulting structure is called a **stratified hierarchy** of partitions. Conversely, every stratified hierarchy of partitions determines a pseudo-ultrametric on the domain. For further details, the reader is referred to [5].

In our case, if an energy partition $\Pi(\delta, S)$ determined by a pseudo-ultrametric δ is considered, the previous

properties imply that every influence zone can be expressed as a lower level set of the energy. Moreover, since the partition does not depend on the location of the sources inside the zones, the problem of selecting a set of sources can be addressed in this case through the choice of a radius r . These particular energy partitions will be denoted by $\Pi(\delta, r)$.

6.2. A Contrast Ultrametric

In this paragraph, a ultrametric energy that extracts the global contrast of an image is defined. For this purpose, the bijection with the class of stratified hierarchies is fundamental, as it provides a constructive definition for this type of pseudo-metrics.

Indeed, as a consequence of the properties presented the previous paragraph, the distance between two points $x \in Z_i$ and $y \in Z_j$ in a ultrametric partition $\Pi(\delta, r)$ can be expressed as a *dissimilarity measure* between the zones Z_i and Z_j . In the discrete space, this remark allows to construct the energy through a region merging strategy.

The idea of progressively merging regions of an initial partition has been used since the early days of computer vision to address segmentation problems [7, 17]. In general, this type of methods, often called *bottom-up* approaches, can be implemented efficiently using a region adjacency graph (RAG), as described in [13, 41]. A RAG is an undirected graph where the nodes correspond to connected regions of the domain. The links encode the vicinity relation and are weighted by the dissimilarity.

Therefore, in this context, the choice of an initial partition and the definition of a dissimilarity measure determine an order for the merging. Then, removing the links of the RAG for increasing values of the dissimilarity and merging the corresponding regions produces a family of nested partitions.

Typically, the dissimilarity expresses a notion of resemblance between neighboring regions and many examples have been proposed in the vast literature on the subject. A simple case is the difference of the average gray level in the regions [9, 22, 32], first used in the classical region growing algorithm [7] and noted by d^a . However, many authors take also into account factors like the variance and the size of the regions [4, 33], the orientation and the texture [42].

Nevertheless, it should be noted that an arbitrary dissimilarity does not necessarily coincide with a ultrametric. In order to produce a stratified hierarchy of

partitions, the dissimilarity must be increasing with the order of merging. Note that this condition is seldom satisfied by the examples found in the literature. When the dissimilarity is not increasing, a stratification index for the hierarchy of partitions can still be defined by considering any increasing function of the merging order. However, in this case, the resulting ultrametric is no longer directly related to the dissimilarity.

The goal in this paragraph was to construct an ultrametric expressing the global contrast of the original image. A natural candidate is the dissimilarity d^a . This option suffers nonetheless from two drawbacks: first, d^a is not increasing and, second, since its definition uses all the information in the zones, its value may not reflect the real contrast. Consequently, a merging process governed by this dissimilarity can create artificial contours when the gray level inside the regions varies gradually.

In order to address these issues, in our case the dissimilarity was constructed using only boundary information and was measured directly on the initial partition. For the examples presented in this paper, the dis-

similarity, noted by d^c , was defined as the average gray level difference in the common boundary of the zones, measured in the extrema mosaic. As a consequence of this choice, d^c is increasing with the merging order and the corresponding pseudo-ultrametric, noted by δ^c , is strongly related to the contrast information provided by the original image.

Figure 7 illustrates the application of our contrast measure on a test image. Top-left shows the initial partition, the extrema mosaic of the *cameraman*. The other images present three energy partitions for increasing radii on the ultrametric balls: $\Pi(\delta^c, 34)$ on top-right, $\Pi(\delta^c, 69)$ on bottom-left and $\Pi(\delta^c, 150)$ on bottom-right; the maximal value of the stratification index in this example was 211. It can be observed how the ultrametric extracts effectively the contrast information in the image: as the radius is increased, only the most contrasted regions remain in the partition, regardless of their size or homogeneity.

A classical example of ultrametric in mathematical morphology is derived from the construction of the watershed transformation [6]. Intuitively, this method can

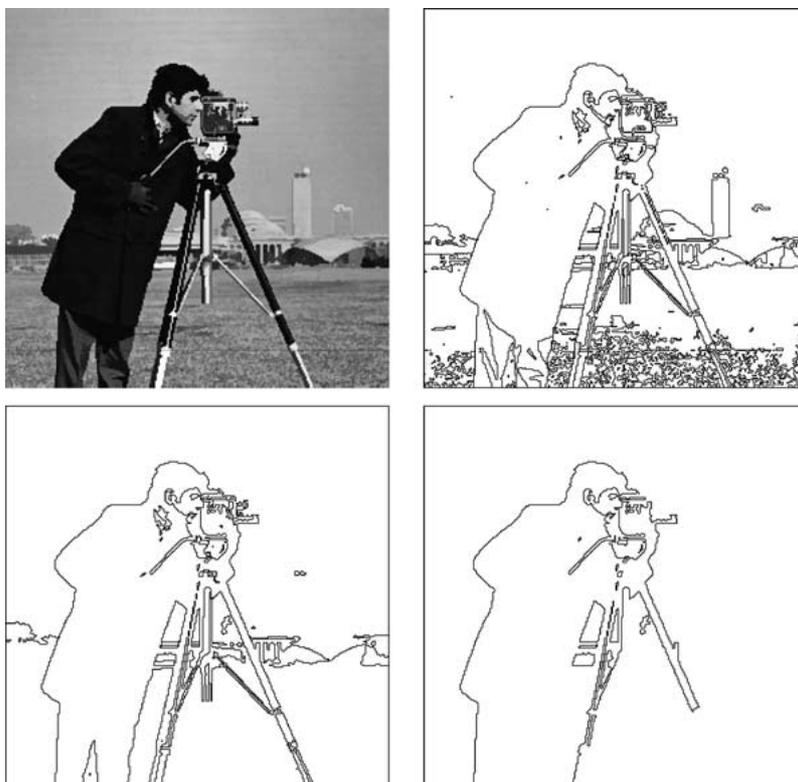


Figure 7. Energy partitions associated to the contrast ultrametric δ^c (see text).

be summarized as follows: an image, considered as a topographical surface, is flooded from its regional minima. The water forms lakes in the valleys and, when two lakes meet, they are merged. Thus, increasing levels of water produce coarser partitions. When the image is a gradient's modulus, the resulting hierarchy is known as the *dynamics* [14]. In terms of a region merging process, the initial partition is composed by the catchment basins of the minima and the dissimilarity is defined as the height of the lowest pass point between two adjacent lakes, i.e., the minimal value of the gradient in the common border of the regions [30]. Therefore, the dynamics hierarchy also induces an ultrametric. However, since its definition is based on a gradient image, the result depends on the choice of a discrete approximation for the gradient. Additionally, the linear smoothing performed by most gradient operators in order to well-posed differentiation, may produce a loss of resolution for this method.

The application of our approach to edge detection, as well as a comparison with the continuous definition of the watershed transform can be found in [1].

7. Derived Ultrametrics

The previous section presented the application of the ultrametric δ^c to the extraction of the contrast of an image. This information is often useful to describe the image structure, as shown on the left column of Fig. 8, where the energy partition $\Pi(\delta^c, 75)$ and its mosaic image are displayed, for the example in Fig. 5. However,

contrast is just one among the factors taken into account in high-level vision tasks. As an example, the central column of Fig. 8 shows a critical index for the contrast ultrametric: on top, the partition $\Pi(\delta^c, 128)$ can be observed and, on bottom, $\Pi(\delta^c, 129)$. At this point, the wings of the butterfly merge with the background. Since the contrast between the light patterns and the dark wings is greater than the contrast between the wings and the background, none of the zones determined by the partitions $\Pi(\delta^c, \lambda)$ corresponds to the entire butterfly (wings and patterns).

A hierarchy of partitions is an interesting representation of the image for several applications like content based image retrieval and object oriented coding. In such applications, the usefulness of the representation is measured by its ability to provide a meaningful partition for any required number of regions. In order to use our approach in those situations, considering only the contrast information may prove restrictive. This fact is illustrated on top-left of Fig. 8, where the partition $\Pi(\delta^c, 203)$ is displayed. If a description with two regions of this image is required, it seems clear that the contrast is insufficient.

In this section, we discuss the definition of ultrametrics whose energy partitions represent more accurately the image structure, when a small number of regions is required. The bottom-right image of Fig. 8 shows an example of the energy partitions that will be constructed.

For this purpose, other perceptually important characteristics of the zones are taken into consideration,

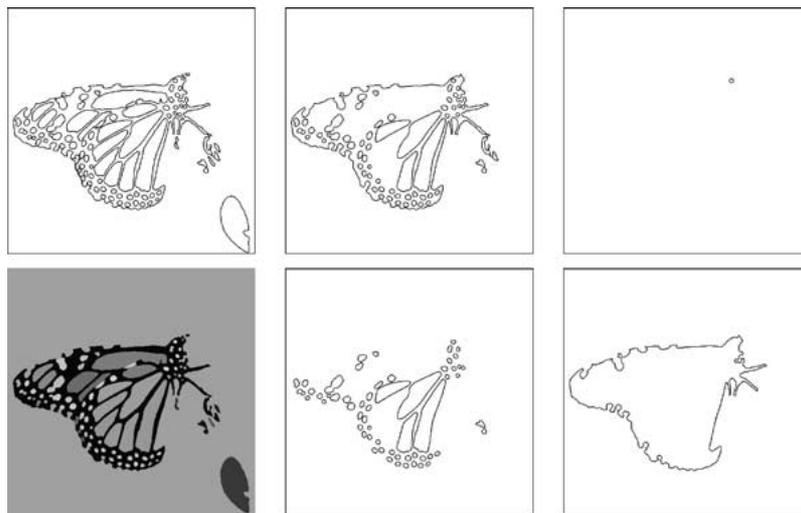


Figure 8. Limitations of using only the contrast information (see text).

in order to complement the boundary information supplied by the contrast. Precisely, an **attribute**, a positive real valued function \mathcal{A} , is defined for every zone. The attribute is required to be increasing with the inclusion order:

$$r_1 \subset r_2 \Rightarrow \mathcal{A}(r_1) \leq \mathcal{A}(r_2),$$

for all connected regions of the domain r_1 and r_2 . In general, the attribute can be calculated using the inter-

nal properties of the zone; the simplest example of an increasing attribute is the area of the zone.

Then, starting from d^c , the dissimilarity associated to δ^c , a new dissimilarity d' can be defined by the formula:

$$d'(Z_1, Z_2) = d^c(Z_1, Z_2) \cdot \min\{\mathcal{A}(Z_1), \mathcal{A}(Z_2)\}.$$

Since \mathcal{A} and d^c are increasing, d' induces a ultrametric. This new energy takes into account the internal as well as the boundary information. For the examples



Figure 9. Top: original images. Middle: energy partitions. Bottom: mosaic images.

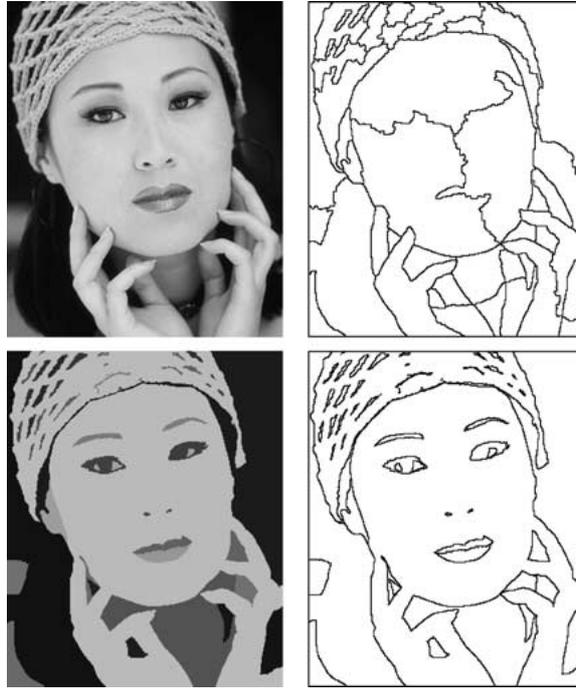


Figure 10. Comparison with the volume hierarchy (see text).

presented in this paper, the attribute was defined as:

$$\mathcal{A}(Z) = A(Z)^\alpha,$$

where A denotes the area of the zone Z and the parameter $\alpha \geq 0$ weights the balance between contrast and area. Thus, the choice of α can be seen as the introduction of higher-level information, allowing the ultrametric to adapt to the image content. The result on bottom-right of Fig. 8 was obtained with $\alpha = 1$.

Figure 9 presents two examples of segmentations obtained with this method. The top row displays the original images, the central row shows the energy partitions and, on bottom, the corresponding mosaic images, with average gray level as the zone model, can be observed. In these examples, the number of regions was 35 and the weight α was set to 1, so the attribute is the area of the zone. Note how, in spite of the low number of regions and the simplicity of the elements considered for the definition of the ultrametric, the main features of the scenes are reconstructed in the segmentations.

The approach presented in this section can be seen as an extension of the *flooding hierarchies* used in morphological segmentation [29, 31, 40]. In this method, during the watershed flooding of a gradient image, the

depth, the area or the volume of the lakes is measured in order to construct hierarchies of partitions.

Figure 10 compares our method with the flooding hierarchy determined by the volume. The top-right image shows the result obtained by selecting the 50 lakes with biggest volume in the morphological gradient of the image on top-left. The result is unsatisfactory because large and shallow lakes have bigger volume than small and deep ones. In contrast, the bottom row shows the mosaic image and the energy partition obtained with our method for the same number of zones, with $\alpha = 0.15$. This example shows the interest in the introduction of the weight α , in order to allow greater versatility for the ultrametric to adapt to the type of image.

8. Conclusion

The present paper discussed the application of energy partitions to image segmentation. The approach consisted in the formulation of the problem as the tessellation of the image domain determined by an energy and a set of sources. The study of two types of energies allowed the interaction with the image at different levels of analysis.

The examples of energies presented in this paper can extract the geometric structure of an image when the objects represented are relatively homogeneous. However, in order to treat highly textured or noisy images with these energies, a pre-processing step should be considered. Alternatively, we are presently working on the inclusion of the texture information in the definition of the energy.

Finally, though this paper focused on real-valued images, a straightforward extension of the method to color images can be obtained by considering only their lightness channel. Additionally, present work also includes the generalization of our approach to vector-valued images.

Acknowledgments

The first author wishes to thank P. Monasse and J.M. Morel for introducing him to the inspiring work of A.S. Kronrod.

The authors also thank the reviewers for their valuable comments and suggestions to improve the quality of this paper.

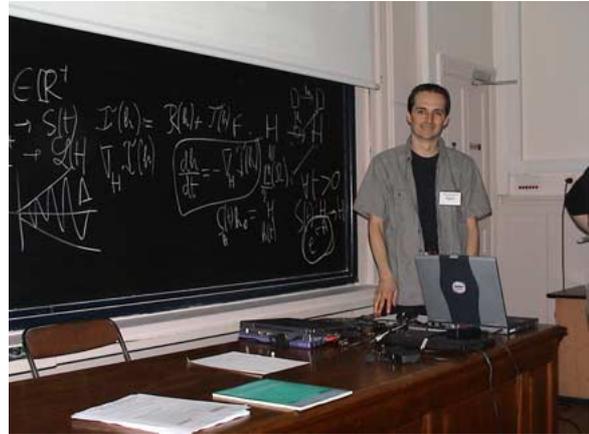
References

1. P.A. Arbeláez and L.D. Cohen, "The extrema edges," in *Proc. Scale-Space'03*, Skye, UK, 2003, pp. 180–195.
2. P.A. Arbeláez and L.D. Cohen, "Path variation and image segmentation," in *Proc. EMMCVPR'03*, Lisbon, Portugal, 2003, pp. 246–260.
3. F. Aurenhammer and R. Klein, *Handbook of Computational Geometry*, Chapt. 5: Voronoi Diagrams, Elsevier Science Publishing, 2000, pp. 201–290.
4. J. Beaulieu and M. Goldberg, "Hierarchy in picture segmentation: A stepwise optimization approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 2, pp. 150–163, 1989.
5. J.P. Benzécri, *L'Analyse des Données. Tome I: La Taxinomie*, 4th edition, Dunod: Paris, 1984.
6. S. Beucher and F. Meyer, *Mathematical Morphology in Image Processing*, Chapt. 12: The Morphological Approach to Segmentation: The Watershed Transformation, Marcel Dekker, 1992, pp. 433–481.
7. C.R. Brice and C.L. Fenema, "Scene analysis using regions," *Artificial Intelligence*, Vol. 1, pp. 205–226, 1970.
8. L.D. Cohen and R. Kimmel, "Global minimum for active contour models: A minimal path approach," *International Journal of Computer Vision*, Vol. 24, No. 1, pp. 57–78, 1997.
9. L.D. Cohen, L. Vinet, P. Sander, and A. Galalowicz, "Hierarchical region based stereo matching," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1989.
10. L.D. Cohen, "Multiple contour finding and perceptual grouping using minimal paths," *Journal of Mathematical Imaging and Vision*, Vol. 14, No. 3, pp. 225–236, 2001.
11. T. Deschamps and L.D. Cohen, "Fast extraction of minimal paths in 3D images and applications to virtual endoscopy," *Medical Image Analysis*, Vol. 5, No. 4, pp. 281–299, 2001.
12. E.W. Dijkstra, "A Note on two problems in connection with graphs," *Numerische Mathematic*, Vol. 1, pp. 269–271, 1959.
13. L. Garrido, P. Salembier, and D. Garcia, "Extensive operators in partition lattices for image sequence analysis," *Signal Processing*, Vol. 66, No. 2, pp. 157–180, 1998. Special Issue on Video Sequence Segmentation.
14. M. Grimaud, "New measure of contrast: Dynamics," in *Image Algebra and Morphological Processing III*, San Diego, USA, 1992.
15. M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Birkhauser: Boston, 1999.
16. E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer Verlag, 1969.
17. S.L. Horowitz and T. Pavlidis, "Picture segmentation by a directed split-and-merge procedure," in *Proceedings of the Second International Joint Conference on Pattern Recognition*, 1974, pp. 424–433.
18. C. Jordan, "Sur la Série de Fourier," *Comptes Rendus de l'Académie des Sciences. Série Mathématique.*, Vol. 92, No. 5, pp. 228–230, 1881.
19. J. L. Kelley, *General Topology*, Springer, 1975.
20. R. Kimmel and A.M. Bruckstein, "Global shape from shading," *Computer Vision and Image Understanding*, Vol. 62, No. 3, pp. 360–369, 1995.
21. R. Kimmel, N. Kiryati, and A.M. Bruckstein, "Distance maps and weighted distance transforms," *Journal of Mathematical Imaging and Vision*, Vol. 6, pp. 223–233, 1996. Special Issue on Topology and Geometry in Computer Vision.
22. G. Koepfler, C. Lopez, and J.M. Morel, "A multiscale algorithm for image segmentation by variational method," *SIAM Journal on Numerical Analysis*, Vol. 31, No. 1, pp. 282–299, 1994.
23. A.S. Kronrod, "On Functions of two variables," *Uspehi Mathematical Sciences*, Vol. 5, No. 35, (In Russian), 1950.
24. R. Kruse and A. Ryba, *Data Structures and Program Design in C++*. Prentice Hall: New York, 1999.
25. R. Malladi and J. Sethian, "A unified approach to noise removal, image-enhancement, and shape recovery," *IEEE Transactions on Image Processing*, Vol. 5, No. 11, pp. 1554–1568, 1996.
26. P. Maragos and M.A. Butt, "Curve evolution, differential morphology and distance transforms applied to multiscale and eikonal problems," *Fundamenta Informaticae*, Vol. 41, pp. 91–129, 2000.
27. D. Marr, *Vision*, Freeman: San Francisco, 1982.
28. D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in: *Proc. ICCV'01*, Vol. II. Vancouver, Canada, 2001, pp. 416–423.
29. F. Meyer, A. Oliveras, P. Salembier, and C. Vachier, "Morphological tools for segmentation: Connected filters and watersheds," *Annals of Telecommunications*, Vol. 52, No. 7/8, pp. 367–379, 1997.
30. F. Meyer, "Morphological segmentation on a neighborhood graph," *Acta Stereologica*, Vol. 16, No. 3, pp. 175–182, 1997.
31. F. Meyer, "Hierarchies of partitions and morphological segmentation," in: *Scale Space and Morphology in Computer Vision*, M. Kerckhove (ed.), 2001, pp. 161–182.

32. A. Montanvert, P. Meer, and A. Rosenfeld, "Hierarchical image analysis using irregular tessellations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 4, pp. 307–316, 1991.
33. P. Nacken, "Image segmentation by connectivity preserving re-linking in hierarchical graph structures," *PR*, Vol. 28, No. 6, pp. 907–920, 1995.
34. I.P. Natanson, *Theory of Functions of a Real Variable*. New York: Frederick Ungar Publishing, 1964.
35. A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd edition, Wiley, 2002.
36. S. Osher and L. Rudin, "Feature-oriented image enhancement using shock filters," *NumAnal*, Vol. 27, No. 4, pp. 919–940, 1990.
37. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, Vol. 60, pp. 259–268, 1992.
38. J. Serra and P. Salembier, "Connected operators and pyramids," in: *Image Algebra and Mathematical Morphology*, SPIE (Ed.), Vol. 2030. San Diego CA., 1993, pp. 65–76.
39. J.A. Sethian, *Level Set Methods and Fast Marching Methods*, 2nd edition Cambridge University Press: Cambridge, UK, 1999.
40. C. Vachier, "Extraction de caractéristiques, segmentation d'Image et morphologie mathématique," Ph.D. thesis, Ecole des Mines de Paris, 1995.
41. T. Vlachos and A.G. Constantinides, "Graph-theoretical approach to colour picture segmentation and contour classification," in: *IEE Proc. Vision, Image and Sig. Proc.*, Vol. 140, pp. 36–45, 1993.
42. W. Yu, J. Fritts, and F. Sun, "A hierarchical image segmentation algorithm," in: *Proc. ICME'02*, 2002, pp. 221–224.



Pablo Andrés Arbeláez finished his B.Sc. studies in pure mathematics at the Universidad de los Andes in Bogotá, Colombia in 1997. In 2000, he received his M.Sc. degree from the Université Paris Dauphine, France in applied mathematics. Since then, he has been preparing his Ph.D. in image analysis and computer vision at the CEREMADE in Paris, France.



Laurent D. Cohen was at Ecole Normale Supérieure Ulm in Paris from 1981 to 1985. He received Master's and Ph.D. degrees in Applied Mathematics from Paris 6 in 1983 and 1986. From 1985 to 1987, he was member at the Computer Graphics and Image Processing group at Schlumberger Palo Alto Research, California and Schlumberger Montrouge Research, and remained consultant there for a few years afterwards. He began working with INRIA, France in 1988, mainly with the medical image understanding group Epidaur. Since 1990, he is Research Scholar (Charge then Directeur de Recherche) with CNRS in the Applied Mathematics and Image Processing group at CEREMADE, University Paris-Dauphine. His research interests and teaching at the university are applications of variational methods and Partial Differential Equations to Image Processing and Computer Vision, like deformable models, minimal paths, surface reconstruction, Image registration, Image segmentation and restoration. He obtained CS 2002 Prize for Image and Signal Processing. He has been member in program committees for boards for about 20 international conferences.