HAL
open science

Global-local optimizations by hierarchical cuts and climbing energies

Bangalore Ravi Kiran, Jean Serra

To cite this version:

Bangalore Ravi Kiran, Jean Serra. Global-local optimizations by hierarchical cuts and climbing energies. Pattern Recognition Letters, 2014, 47 (1), pp.12-24. 10.1016/j.patcog.2013.05.012 . hal-00802978v2

HAL Id: hal-00802978
https://hal.science/hal-00802978v2

Submitted on 29 Apr 2013

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.

Global-local optimizations by hierarchical cuts and climbing energies

B. Ravi Kiran, Jean Serra

Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, A3SI, ESIEE
{j.serra, kiranr,} @esiee.fr

Abstract

Hierarchical segmentation is a multi-scale analysis of an image and provides a series of simplifying nested partitions. Such a hierarchy is rarely an end by itself and requires external criteria or heuristics to solve problems of image segmentation, texture extraction and semantic image labelling. In this theoretical paper we introduce a novel framework: hierarchical cuts, to formulate optimization problems on hierarchies of segmentations. Second we provide the three important notions of *h-increasing*, *singular*, and *scale increasing* energies, necessary to solve the global combinatorial optimization problem of partition selection and which results in linear time dynamic programs. Common families of such energies are summarized, and also a method to generate new ones is described. Finally we demonstrate the application of this framework on problems of image segmentation and texture enhancement.

Keywords: Hierarchical segmentation, Climbing optimization, Mathematical Morphology, Energy minimization, Dynamic programming

1. Introduction

To segment an image by a global constraint classically means to associate a numerical energy with every possible partition of the space where this image is defined. The best partition is then that which minimizes the energy. Is this meaningful ? Let us suppose that the energies range from 0 to 10^3. Using the formula for the classical Bell's number, a digital square of 5×5 pixels has 4.6×10^{18} different partitions possible [9]. Each value of energy thus maps on to millions of billions (4.6×10^{15}) of partitions. What do we minimize

here? Which implicit assumptions underlie the methods which give a unique minimal cut?

There are only two ways for obtaining (or hoping) uniqueness: by limiting the number of partitions, and by imposing constraints to the energy. To limit the number of partitions, we can think of cuts on hierarchies, which provide strong restrictions. To constrain the energy, we can try and replace the lattice of the integers by another one, more comprehensive, e.g. a lattice of partitions, and make hold the minimizations on it. But how to create a lattice of partitions from a given energy? Which conditions must we introduce? And if uniqueness is finally ensured (the lattice structure is precisely made for that), how to reach the minimal cut in the maze of all partitions? By means of which vital thread?

There have been several approaches to global constraints for optimization. There are two methods we contrast here: First, the graph cuts based optimization, popularized by Y. Boykov [7], second, partition selection from hierarchies of partitions. The former emphasize the use of seeds, in addition, they view the space as a one scale structure. This perspective is illustrated by the search for a maximum flow in a directed graph, whose segmentation applications include the optimization of conditional random field (CRF) [22]. The latter approaches emphasize the scaling of the space by means of hierarchies, and attach less importance to labelling questions, in a first step at least.

This paper focusses on the second type of global constraints, which are approached from the viewpoint of hierarchical cuts (h-cuts) theory[1]. A hierarchy, or pyramid, of image segmentations is understood as a series of progressive simplified versions of an initial image, which result in increasing partitions of the space. How can these partitions cooperate and summarize the hierarchy into a unique cut, optimal in some sense. Three questions arise here, namely:

1. Given a hierarchy H of partitions and an energy ω on the partial partitions, how to combine the classes of this hierarchy for obtaining a new partition that minimizes ω, and which can be determined easily?
2. When one energy ω depends on an integer j, i.e. $\omega = \omega^j$, how to generate a sequence of optimal partitions that increase with j, which

[1]This work received funding from the Agence Nationale de la Recherche through contract ANR-2010-BLAN-0205-03 KIDIKO.

therefore should form a optimal hierarchy?

3. Most of the segmentations involve several features (colour, shape, size, etc.) that we can handle with different energies. How to combine them?

These questions have been taken up by several authors, over many years, and by various methods. The most popular energies ω for hierarchical partitions derive from that of D. Mumford and J. Shah [25], in which a data fidelity term is summed up with a boundary regularization term. The optimization turns out to be a trade-off between these two constraints. In [20], for example, G. Koepfler, C. Lopez, and J.M. Morel build a pyramid of segmentations, from fine to coarse, by progressively giving more and more weight to the second term of Mumford and Shah functional. They stop the region growing when a certain number of regions is reached. The method initiated by P. Salembier and L. Garrido for generating thumbnails rests on the same type of energy [33]. They interpret the optimal cut as the most accurate image simplification for a given compression rate. The approach has been extended to additive energies by L. Guigues et Al. [17]. It is always assumed, in all these studies, that the energy of any partial partition equals the sum of the energies of its classes, which considerably simplifies the combinatorial complexity, and answers the above two questions 1 and 2.

However, one can wonder whether additivity is the very underlying cause of the simplifications, since P. Soille's constrained connectivity [38], where the addition is replaced by the supremum, satisfies similar properties. Finally, one finds in literature a third type of energy, which holds on nodes only, and no longer on partial partitions. It appears in the method for labeling of P. Arbeláez [3], or in the studies of H.G. Akçay and S. Aksoy, in [1]. And again, it yields optimal cuts.

Is there a common denominator to all these approaches, more comprehensive than just additivity, and which explains why they always lead to unique optima? The following paper is a theoretical attempt to delimit this central concept, and to give answers to the above questions from 1) to 3). The theory is established in section 3 to 5, section 6 presents the algorithms, which are then applied to the two main families of climbing energies in section 7. Before the conclusion, the approach is extended to partial optimization in section 9, and some bridges between graph cuts and hierarchical optimizations are given in section 10.

2. Basic notions: Hierarchies and partitions

This section provides the background required to understand this paper. The usual distinction between continuous and digital spaces is not appropriate for the general theory developed in sections 3 to 6. What is actually needed reduces to the two following hypotheses, which are assumed over the whole paper:

1. the space E to partition is topological,
2. the smallest partition π_0 of E has a finite number of classes.

The first assumption allows us to speak of frontiers between classes, or edges. The second one aims to avoid fractalities, and to permit various inductions, in Proposition 3.2 and in Algorithm 1, among others. Some additional hypotheses are introduced when the energies are particularized in section 7, e.g. "the classes are connected sets", or "the edges are simple arcs of $\mathbb{R}^2$". None of these assumptions are specific to image analysis. Space E may be the concern of parameters, semantic entities, grammars, NASDAQ quotations, or chamber music as well.

2.1. Partitions, partial partitions

Intuitively, a partition of E of the space under study (Euclidean, digital, graph, or else) is a division of E into regions that do not overlap, and whose union restores E in its entirety. These regions are called classes. More formally, one obtains the classes of a partition by means of an extensive mapping $S : E \to \mathcal{P}(E)$ such that,

$$x, y \in E \quad \Rightarrow \quad S(x) = S(y) \;\text{ or }\; S(x) \cap S(y) = \varnothing.$$

Below, the symbols S, T stand for classes, and π for partitions. Partition π_1 is smaller than partition π_2 when each class of π_1 is included in a class of π_2. This condition provides an ordering on the partitions, called refinement, which in turn induces a complete lattice.

Following Ch. Ronse [30], a partition $\pi(S)$ associated with a set $S \in \mathcal{P}(E)$ is called *partial partition (in short p.p.)* of E of support S. In particular, the partial partition of S into the single class S is denoted by $\{S\}$. The family of all partial partitions of set E is denoted by $\mathcal{D}(E)$, or simply by $\mathcal{D}$.

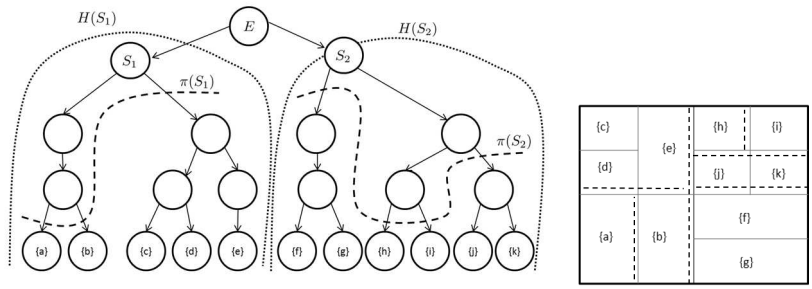

Figure 1: Hierarchy representation using a dendrogram and corresponding leaves partition shown in full lines, while the union of cuts $\pi(S_1) \cup \pi(S_2)$ in dotted lines.

2.2. Hierarchies of partitions

A hierarchy H is a chain of partitions π_i, i.e.

$$H = \{\pi_i, 0 \leq i \leq n \mid i \leq k \leq n \Rightarrow \pi_i \leq \pi_k\}, \tag{1}$$

where π_0 is the finest partition and π_n is the partition $\{E\}$ of E in a single class. The classes of π_0 are called the *leaves*, and E is the *root*. Since the number of leaves of π_0 is finite (as we have assumed above), the number n of different partitions of H is also finite [2]. The intermediary classes are called *nodes*. If the q classes of the partition $\pi(S)$ are $\{T_u, 1 \leq u \leq q\}$, one writes

$$\pi(S) = T_1 \sqcup ..T_u.. \sqcup T_q,$$

where the symbol $\sqcup$ indicates that the classes are concatenated. Given two p.p. $\pi(S_1)$ and $\pi(S_2)$ having disjoint supports, $\pi(S_1) \sqcup \pi(S_2)$ is the p.p. whose classes are either those of $\pi(S_1)$ or those of $\pi(S_2)$.

Let $S_i(x)$ be the class of partition π_i of H at point $x \in E$. Expression (1) means that at each point $x \in E$ the family $\{S_i(x), x \in E, 0 \leq i \leq n\}$ of those classes $S_i(x)$ that contain x forms a finite chain of nested sets from the leaf $S_0(x)$ to E.

$$\mathcal{S}(x) = \{S_i(x), 0 \leq i \leq n\}. \tag{2}$$

[2]One could argue that all components being finite, the underlying space E does not need to be infinite. However, some problems require a finite number of leaves embedded in a continuous space, e.g. ground truth by distance function [21].

Conversely, according to a classical result [8], a family $\{S_i(x), x \in E, 0 \leq i \leq n\}$ of indexed sets generates the classes of a hierarchy iff, for $i \leq j$ and $x, y \in E$

$$S_i(x) \subseteq S_j(y) \text{ or } S_i(x) \cap S_j(y) = \varnothing, \tag{3}$$

conditions which mean that the classes form an ultra-metric space [4], [23]. The partitions of a hierarchy may be represented by their classes, via a dendrogram, i.e. a tree where each node of bifurcation is a class S, or by their frontiers, via the saliency map of the edges, which indicates the level in the hierarchy when an edge disappears [27],[14]. The first representation is depicted in Figure 1, the second one in Figure 2. The classes of π_{i-1} at level $i-1$ which are included in class S_i of level i are said to be *the sons* of S_i. Clearly, the descendants of each node S form in turn a hierarchy $H(S)$ of root S, which is included in the complete hierarchy $H = H(E)$. One denotes by $\mathcal{S}(E)$, or just $\mathcal{S}$, the set of all classes S of all partitions involved in H.

The hierarchy can be loosely seen as a set of partitions containing superpixels of increasing sizes. Here we don't use the superpixel terminology, and prefer to distinguish between the class and partial partition.

Figure 2: Ducks: Initial RGB image, Luminance of RGB image, Saliency of partitions from Luminance (inverted to see with better contrast)

2.3. Generating hierarchies of segmentations

In the paper, the focus is not on the methods for obtaining hierarchies of segmentations, they are considered as inputs. The main techniques for hierarchical segmentation include the various Matheron semi-groups of connected filters (openings, alternating sequential filters) [32], the progressive floodings of watersheds [24], [14] and [3], the hierarchies obtained by increasing connections [2], [38], and the functional minimizations of Mumford and Shah type [25]. In addition, the learning strategies for segmentation, as developed by [4], among others, lead to very significant watersheds hierarchies. One can

also quote the approach in [10] where additive functionals are enriched by the introduction of shape descriptors, which yield compact representations in [6].

2.4. Cuts in a hierarchy

Any partition π of E whose classes are taken in $\mathcal{S}$ defines a *cut* π in a hierarchy H. The set of all cuts of E is denoted by $\Pi(E) = \Pi$. Every "horizontal" section $\pi_i(H)$ at level i is obviously a cut, but several levels can cooperate in a same cut, such as $\pi(S_1)$ and $\pi(S_2)$, drawn with thick dotted lines in Figure 1, where the partition $\pi(S_1) \sqcup \pi(S_2)$ generates a cut of $H(E)$. One can also define cuts inside any sub-hierarchy $H(S)$ of summit S, and similarly $\Pi(S)$ stands for the family of all cuts of $H(S)$.

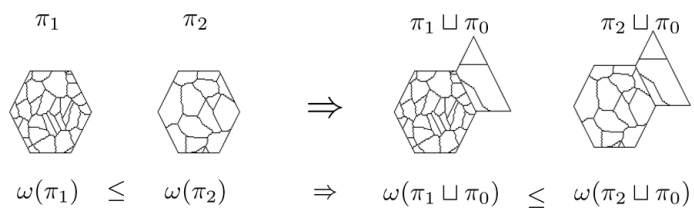

Figure 3: $\pi_1(S)$ and $\pi_2(s)$ are partial partitions on hexagonal support S, h-increasingness

In particular, the refinement supremum $\pi = \pi_1 \vee \pi_2$ of two cuts π_1 and π_2 is the cut where the class $S(x)$ at each leaf x is $S(x) = S_1(x) \cup S_2(x) = S_1(x)$ or $S_2(x)$.

3. Cuts and energies

3.1. Energies

An energy on the set $\mathcal{D}$ of all p.p. of E is a numerical function $\omega : \mathcal{D} \to [0, \infty]$. In the following, $\mathcal{D}$ will be provided with several energies ω, which may satisfy the following axioms:

i) ω is *singular* when the energy $\omega(\{S\})$ of class S is strictly smaller than, or larger than, any p.p. of S, i.e.

$$\omega(\{S\}) > \vee\{\omega(\pi(S))\} \text{ or } \omega(\{S\}) < \wedge\{\omega(\pi(S))\}, \pi(S) \text{ p.p. of } S. \quad (4)$$

For example, in $\mathbb{R}^2$ or $\mathbb{Z}^2$, the number of classes of a p.p., or the sum of their perimeters, or again the sup of their areas, are singular energies, but the sum of their areas is not.

ii) ω is *h-increasing*, i.e.

$$\omega(\pi_1) \leq \omega(\pi_2) \quad \Rightarrow \quad \omega(\pi_1 \sqcup \pi_0) \leq \ \omega(\pi_2 \sqcup \pi_0), \tag{5}$$

where π_1 and π_2 are two p.p. of same support S, and π_0 a p.p. of support S_0 disjoint of S [36]. The geometrical meaning of Rel.(5) is shown in Figure 3, and several examples are proposed in sections 4.2 and 7.

The first property results in the energetic lattice of cuts in (3.2) and the second results in the climbing property in (4.3). We need both indeed, which leads us to the following definition:

Definition An energy ω which is both h-increasing and singular is said to be climbing.

3.2. Energetic orderings of the cuts

When an energy ω is allocated to the p.p. of E, the cuts that minimize ω provide optimal segmentations of E relative to ω. However, one can wonder about the meaning of such best cuts. The family of possible cuts being finite, and the energy ω being a positive number, we are sure to always find a cut that minimizes ω. And not only one but billions of billions, as we already saw.... The problem is ill-posed because the minimization holds on the lattice of the positive numbers. We must change the approach, well-pose the problem by equipping the set of all cuts of H with an ordering which involves ω, and make the minimization directly hold on the cuts. If we are able to associate a lattice with such an ordering, the uniqueness of the solution will then be ensured.

Theorem 3.1. *Given a hierarchy H, an energy ω induces an energetic ordering on the set $\Pi(E)$ of all cuts of H, if and only if ω is singular. In this ordering, cut $\pi \in \Pi(E)$ is less energetic than cut $\pi' \in \Pi(E)$ w.r.t. ω, and one writes $\pi \leq_\omega \pi'$, when in each class S of the supremum by refinement $\pi \vee \pi'$ the p.p. of π inside S has an energy smaller or equal to that of π inside S. Equivalently, for each leaf $x \in E$*

a) either the class $S(x)$ of π is the support of a p.p. χ' of π' and $\omega(\{S\} \leq \omega(\chi')$,

b) or the class $S'(x)$ of π' is the support of a p.p. χ of π and $\omega(\chi) \leq \omega(\{S'\})$.

Proof. The equivalence of the two formulations is a consequence of Rel. 3, which shows that each class of $\pi \vee \pi'$ is either a class of π or of π'. The reflexivity, in statements *a)* and *b)* is obvious. For the transitivity, consider $\pi_1, \pi_2, \pi_3 \in \Pi$, with $\pi_1 \leq_\omega \pi_2$ and $\pi_2 \leq_\omega \pi_3$. At leaf x, their three classes are S_1, S_2, and S_3 respectively. If $S_1 = S_2$ or $S_2 = S_3$, the theorem is locally satisfied. If not, one cannot have $S_1 \cup S_3 \subseteq S_2$. Indeed, if $S_3 \subseteq S_2$, there exists a p.p. χ with $\{S_3\} \sqcup \chi = \{S_2\}$, and the assumption $\pi_2 \leq_\omega \pi_3$ implies, by *a)*, that $\omega(\{S_2\} \leq \omega(\{S_3\} \sqcup \chi)$. If in addition $S_1 \subseteq S_2$, i.e. $\{S_1\} \sqcup \chi' = \{S_2\}$, we see similarly that $\omega(\{S_1\} \sqcup \chi') \leq \omega(\{S_2\})$, which contradicts the singularity. Therefore, the three classes S_1, S_2, and S_3 can be ordered in two ways only, namely

$$\begin{aligned} i) \quad & S_1 \subseteq S_3 \quad \text{and} \quad S_2 \subseteq S_3, \\ ii) \quad & S_3 \subseteq S_1 \quad \text{and} \quad S_2 \subseteq S_1. \end{aligned}$$

In case $i)$, there exist two p.p. ζ and ζ' with $\{S_1\} \sqcup \zeta\} = \{S_3\}$ and $\{S_2\} \sqcup \zeta'\} = \{S_3\}$. As $\pi_2 \leq_\omega \pi_3$, we have, by *a)*, $\omega(\{S_2\} \sqcup \zeta') \leq \omega(\{S_3\})$. Therefore, by singularity, all p.p. of $\{S_3\}$ have energies $\leq \omega(\{S_3\})$. In particular $\omega(\{S_1\} \sqcup \zeta) \leq \omega(\{S_3\})$, which shows that transitivity is fulfilled at leave x. In case $ii)$, a similar proof yields the same conclusion, so that finally $\pi_1 \leq_\omega \pi_3$.

For the anti-symmetry, we must prove that $\pi \leq_\omega \pi'$ and $\pi' \leq_\omega \pi$ imply that $\pi = \pi'$. Suppose that the class $S'(x)$ of π' is the support of a p.p. χ made of more than one class of π. By applying the case *b)* of the theorem to the inequality $\pi \leq_\omega \pi'$, we have $\omega(\chi) \leq \omega(\{S'\})$. But we are also in case *a)* for $\pi' \leq_\omega \pi$, hence $\omega(\chi) \geq \omega(\{S'\})$, which implies the equality of the two members. But this contradicts the singularity of ω, so that S' is partitioned into a unique class of π, namely S. If we reverse the roles of π and π', we obtain the same result, which is also independent of the choice of the leave x in E. This achieves the proof of anti-symmetry.

Conversely, consider an ordering $\leq_\omega$whose energy would be non singular, and two cuts π and π' identical everywhere except in the class $S'(x)$ of π, where π is locally the p.p. χ. Supposed that $\omega(\chi) = \omega(S'(x))$. This implies $\pi \leq_\omega \pi'$ and also $\pi' \leq_\omega \pi$. However we do not have $\pi' = \pi$ since $\chi \neq S'(x)$. Thus singularity is needed, which achieves the proof. □

A lattice structure directly derives from the ordering $\leq_\omega$:

Proposition 3.2. *The set $\Pi(E)$ of all cuts of $H(E)$ forms a lattice for the energetic ordering $\leq_\omega$. Given a family $\{\pi^j, 1 \leq j \leq p\}$ in $\Pi(E)$, the infimum $\wedge_\omega \pi^j$ (resp. supremum $\vee_\omega \pi^j$) is obtained by taking the partial partitions of lowest energy (resp. highest energy) in each class of the refinement supremum $\vee \pi^j$.*

Proof. Consider two partitions π^1 and π^2. In each class S of the refinement supremum $\pi^1 \vee \pi^2$ one of the two partitions at least is $\{S\}$, and the other a finer p.p. of support S. By singularity of ω they have different energies, so that one can always choose the less (resp. most) energetic one. By doing the same for all classes of $\pi^1 \vee \pi^2$ we obtain the unique largest lower-bound $\pi^1 \wedge_\omega \pi^2$. (resp. smallest upper bound $\pi^1 \vee_\omega \pi^2$) of π^1 and π^2.

Under iteration, the proof extends to any integer n, which achieves it. □

Theorem 3.1 and proposition 3.2 enrich the structure for optimization by providing local interpretation of a global energy , since energetic extrema are now associated with *each class* of $\vee \pi^j$. Note also that any sub-hierarchy of root a node S of H forms in turn an energetic lattice for $\leq_\omega, \vee_\omega$, and $\wedge_\omega$. The "if" statement of the theorem urges us to find singular energies, and the "only if" one tells us that the described localization can only be reached by singular energies. But of course, there may exist other nice orderings of the cuts, on the base of other types energies as well.

In the notation, we distinguish the refinement lattice from the ω-lattice by using for the former the three symbols $\leq, \vee$, and $\wedge$, without ω subscript.

4. Optimization and hierarchical increasingness

The optimization problem can now be stated more precisely. It consists in finding the minimum element of the energetic lattice $(\vee_\omega, \wedge_\omega)$. Three entities are involved, namely:

- A hierarchy/pyramid H of partitions of E which segment an input image,
- An energy ω, i.e. a non negative numerical function over the family $\mathcal{D}(E)$ of all partial partitions of E,
- An "energetic" function f on E which may be the initial image, or another one, which parametrizes the energy ω.

These three pieces of information are independent, and aim to determine the cuts π^* that minimize ω, i.e. such that $\omega(\pi^*) = \inf\{\omega(\pi) \mid \pi \in \Pi(E)\}$. they are called below the *optimal cuts*.

4.1. Optimal cut characterization

Let $\mathcal{H}$ represent the set of all finite hierarchies of partitions of E.

Definition An energy ω on $\mathcal{D}(E)$ is *weakly h-increasing* when for any hierarchy $H \in \mathcal{H}$, any disjoint nodes S and S_0 of H, and any partition π_0 of S_0, we have

$$\omega(\pi^*) = \inf\{\omega(\pi),\ \pi \in \Pi(S)\} \Rightarrow\ \omega(\pi^* \sqcup \pi_0) \leq \inf\{\omega(\pi \sqcup \pi_0),\ \pi \in \Pi(S)\} \quad (6)$$

where $\Pi(S)$ stands for the finite set of all p.p. of node S involved in hierarchy H.

Clearly, h-increasingness implies weak h-increasingness, i.e. Rel.(5) $\Rightarrow$ Rel.(6). More precisely, Rel.(5) has been weakened just enough to obtain the theorem 4.1 of optimal cut working in both senses. Indeed, we now have

Theorem 4.1. *Let $H \in \mathcal{H}$ be a finite hierarchy, and ω an energy on $\mathcal{D}(E)$, and S be a node of H of sons $T_1..T_p$. If $\pi_1^*, ..\pi_p^*$ are optimal cuts of $T_1..T_p$ respectively, then*

$$\pi_1^* \sqcup \pi_2^* .. \sqcup \pi_p^* \quad (7)$$

is an optimal cut of $\Pi(S)\backslash\{S\}$, for any $H \in \mathcal{H}$ and any $T_1..T_p$ in H, if and only if ω is weakly h-increasing.

Proof. Let us prove that Rel.(6) implies that $\pi_1^* \sqcup \pi_2^* .. \sqcup \pi_p^*$ is an optimal cut of S. We firstly limit ourselves to two classes, T_1 and T_2, say, and consider the energy $\omega(\pi_1^* \sqcup \pi_2^*)$. The weak h-increasingness of ω implies that

$$\omega(\pi_1^* \sqcup \pi_2^*) \leq \inf\{\omega(\pi_1^* \sqcup \pi_2), \pi_2 \in \Pi(T_2)\},$$

and that

$$\omega(\pi_1^* \sqcup \pi_2) \leq \inf\{\omega(\pi_1 \sqcup \pi_2), \pi_1 \in \Pi(T_1)\}.$$

Hence

$$\omega(\pi_1^* \sqcup \pi_2^*) \leq \inf\{\omega(\pi_1 \sqcup \pi_2), \pi_1 \in \Pi(T_1), \pi_2 \in \Pi(T_2)\},$$

which shows that $\omega(\pi_1^* \sqcup \pi_2^*)$ is an optimal cut of $\pi(T_1 \cup T_2)\backslash\{T_1 \cup T_2\}$. Under finite iteration, the property extends to $S = \cup\{T_i, 1 \leq i \leq p\}$. Conversely,

suppose that ω is not weakly h-increasing. It means that one can find a node S and a partial partition $\pi_0 \in S_0$, with both S and S_0 in H and such that

$$\omega(\pi^* \sqcup \pi_0) > \inf\{\omega(\pi \sqcup \pi_0),\ \pi \in \Pi(S)\}. \tag{8}$$

Consider the hierarchy H_0 which derives from H by replacing the partitions of S_0 by π_0, at all levels $\leq$ to that of S_0. In H_0, π_0 is a optimal cut of S_0, although, because of Rel.(8) and because the finite number of elements of $\Pi(S)$, $\omega(\pi^* \sqcup \pi_0)$ is not a optimal cut of $S \cup S_0$. This counter example achieves the proof. □

Corollary 4.2. *When ω is h-increasing but not weakly, then the "only if" part of the theorem is no longer true.*

Presented in a stochastic framework where probabilities are assigned to energies, h-increasingness could be interpreted as a Markov chain property (of order one), since when the optimal energies of the sons of S are known, one does not need the knowledge of the descendants below the sons, to determine whether the energy of S is optimal.

When the ω is both h-increasing and singular, and thus climbing, then Theorem 4.1 leads to the following key consequence

Proposition 4.3. *Let ω be a climbing energy. Then for any $H \in \mathcal{H}$ and any node S of H with p sons $T_1..T_p$ of optimal cuts $\pi_1^*..\pi_p^*$, there exists a unique optimal cut of the sub-hierarchy of root S. It is either the cut $\pi_1^* \sqcup \pi_2^*.. \sqcup \pi_p^*$, or the one class partition $\{S\}$ itself:*

$$\omega(\pi^*(S)) = \min\{\omega(\{S\}), \omega(\pi_1^* \sqcup \pi_2^*.. \sqcup \pi_p^*)\} \tag{9}$$

Proposition 4.3 is essential. It governs the choices of models for energies, and their implementations:

Firstly, the obtained optimal cut $\pi^*(E)$, is indeed globally less energetic than any other cut in H. In addition, locally each class $S \in \pi^*(E)$ is less energetic than any p.p. of S into classes of H, and also less energetic than any p.p. composed of classes of H and containing S. This is a strong property of regional minimum. Thus, this optimization is both local and global.

Secondly, the dynamic programming Rel.(9) allows us to find the optimal cut of H in one ascending pass. The nodes of H above the leaves have to be visited according to an order which respects the inclusions. One then compares the energy of each node with that of the p.p. of its sons, and the

less energetic of the two is kept for continuing the ascending pass, and so on until the top node E is reached (the formal algorithm is described in section 6).

Thirdly, the condition (5) of h-increasingness for an energy being a notion independent of any hierarchy, one can use a different ω for each of the n levels of hierarchy H.

Finally, though Proposition 4.3 is *a fortiori* satisfied when ω is weakly h-increasing, and we deal with h-increasingness. Fortunately so, because it is incomparably easier to check the h-increasingness of an energy than its possible weak h-increasingness. Energies are not alywas h-increasing. For example, the energy ω equal to the number $n(\pi)$ of classes of the p.p. π when $n < 3$, and $0.5n$ when not, is not h-increasing.

4.2. uniqueness of the minimum cut

Are the assumptions of proposition 4.3 compatible with each other? Can an energy be both h-increasing and singular? For answering the question, we must involve the minimum m of all the absolute differences of energies involved in the p. p. of H, i.e.

$$m = \inf\{|\omega(\pi) - \omega(\pi')|\,, \omega(\pi) < \omega(\pi')\} \quad \pi, \pi' \text{ p.p. of } H.$$

As the number of p.p. of H is finite, m is strictly positive. Therefore, one can find a ε such as $0 < \varepsilon < m$, and state the following:

Proposition 4.4. *Let ω be a h-increasing energy over $\mathcal{D}$. Introduce the additional energy ω' for all $\{\pi(S) \in \Pi(S),\ S \in \mathcal{S}\}$*

$$\omega'[\pi(S)] = \varepsilon \text{ when } \pi(S) \neq \{S\} \quad \text{and} \quad \omega'[S] = 0 \text{ when not,}$$

with $0 < \varepsilon < m$. Then the sum $\omega + \omega'$ is climbing and provides a unique optimal cut with each sub-hierarchy $H(S)$, namely

$$\begin{aligned} \pi^*(S) &\text{ when } \omega[\pi^*(S)] \neq \omega[\{S\}] \\ \{S\} &\text{ when } \omega[\pi^*(S)] = \omega[\{S\}]. \end{aligned} \tag{10}$$

Proof. The energy ω' is h-increasing, since in implication (5) the two p.p.$\{S\} \sqcup \pi_0$ and $\pi(S) \sqcup \pi_0$ have always more than one class, so that $\omega'(\{S\} \sqcup \pi_0) = \omega'(\pi(S) \sqcup \pi_0) = \varepsilon$ and Rel.(5) is satisfied. Hence $\omega + \omega'$ is h-increasing. As ω' is smaller than m, the two energies ω and $\omega + \omega'$ yield the same minimum cut $\pi^*(S)$ when $\omega[\pi^*(S)] \neq \omega[\{S\}]$, but when not $(\omega + \omega')(\{S\})$ is less energetic than $(\omega + \omega')(\pi^*(S))$, which achieves the proof. □

Mutatis mutandis, the same proof applies also if one takes systematically $\pi^*(S)$ rather than $\{S\}$ in case of equality, or again, given ω_0, if one takes $\pi^*(S)$ when $\pi^*(S) \leq \omega_0$ and $\{S\}$ when not, etc.

What happens when ω is h-increasing but not singular? The number of the optimal cuts explodes, and to attain uniqueness, one must use a complementary energy ω', which is singular. Then in the ω'-energetic lattice, the family of all ω-optimal cuts has a unique minimum (Proposition 3.2). For finding it, it suffices, in the climbing procedure, to take at each node S the less ω'-energetic p.p. among $\{S\}$ and the $\pi^*(S)$ of Eq.(9). This is indeed exactly what we do in the examples below by choosing $\{S\}$ rather than $\pi^*(S)$ when $\omega(\{S\}) = \omega(\pi^*(S))$. Here ω' is nothing but the number of classes.

4.3. Identifying and generating climbing energies

An easy way to obtain a h-increasing energy consists in defining it, firstly, over all sets $S \in \mathcal{P}(E)$, considered as one class partial partitions $\{S\}$, and then in extending it to all partial partitions by some law of composition. Then, the h-increasingness is introduced by the law of composition, and not by $\omega[\mathcal{P}(E)]$.

The first two modes of composition which come to mind are, of course, addition and supremum. The additive mode was studied by L. Guigues under the name of *separable energies* [16], [17], a context in which he established the Rel.(14) below. All classes S of $\mathcal{S}$ are supposed to be connected. Denote by $\{T_u, 1 \leq u \leq q\}$ the q sons which partition the node S, i.e. $\pi(S) = T_1 \sqcup ..T_u.. \sqcup T_q$. Provide the simply connected sets of $\mathcal{P}(E)$ with an arbitrary energy ω, and extend it from $\mathcal{P}(E)$ to the set $\mathcal{D}(E)$ of all partial partitions by using the sums

$$\omega(\pi(S)) = \omega(T_1 \sqcup ..T_u.. \sqcup T_q) = \sum_1^q \omega(T_u). \tag{11}$$

Just as the sum-generated ones, the $\vee$-generated energies on the partial partitions are defined from an energy ω on $\mathcal{P}(E)$ followed by a law of composition, which is now the supremum.

$$\omega(\pi) = \omega(T_1 \sqcup ... \sqcup T_n) = \vee\{\omega(T_i)\}. \tag{12}$$

Proposition 4.5. *Let E be a set and $\omega : \mathcal{P}(E) \to \mathbb{R}^+$ an arbitrary energy defined on $\mathcal{P}(E)$, and let $\pi \in \mathcal{D}(E)$ be a partial partition of classes*

$\{S_i, 1 \leq i \leq n\}$. *Then the two extensions of* ω *to the partial partitions* $\mathcal{D}(E)$ *by addition (Rel. (11)) and by supremum (Rel. (12)) define* h*-increasing energies.*

Proof. Let $S_{i,0}, S_{i,1}, S_{i,2}$ be the classes of π_0, π_1, π_2 respectively in implication (5), and let ω be an additive energy. We draw from the law of composition (11) that $\omega(\pi_1 \sqcup \pi_0) = \omega(\pi_1) + \omega(\pi_0)$, and $\omega(\pi_2 \sqcup \pi_0) = \omega(\pi_2) + \omega(\pi_0)$, which shows that the additive energy ω is h-increasing. Similarly, if ω is $\vee-$generated, then $\omega(\pi_1 \sqcup \pi_0) = \omega(\pi_1) \vee \omega(\pi_0)$, and $\omega(\pi_2 \sqcup \pi_0) = \omega(\pi_2) \vee \omega(\pi_0)$, which achieves the proof. □

Corollary 4.6. *If* $\{\alpha_j, j \in J\}$ *stands for a family of non negative weights, then the weighed sum* $\sum \alpha_j \omega_j$ *and supremum* $\bigvee \alpha_j \omega_j$ *of* h*-increasing energies* ω_j *turn out to be* $h-$*increasing.*

A number of other laws are compatible with h-increasingness, such as the infimum or multiplication. One can also make ω depend on more than one class, on the proximity of the edges, on another hierarchy, etc.. Examples of composition by addition and supremum are demonstrated in Section 7.

5. Scale increasingness and climbing families of energies

5.1. Scale increasingness

There are energies ω which are parametrized by a scale parameter λ, for example the Mumford-Shah functional. In fact, this sort of notion is less associated to a unique energy ω than to a family $\{\omega_\lambda, \lambda \in \Lambda\}$ of energies depending on a positive value λ. We will describe such a scaling in terms of progressively coarser optimal cuts (for the refinement ordering) with respect to λ. In literature, the idea of acting on the λ coefficient of Mumford and Shah functional in Eq.(17) goes back to G. Koepfler, C. Lopez, and J.M. Morel [20], who used λ for generating the input pyramid H.

The family $\{\omega_\lambda, \lambda \in \Lambda\}$ provides hierarchy H with a sequence of p optimal cuts $\pi^{\lambda *}$, of labels $\lambda \in \Lambda$. *A priori*, the $\pi^{\lambda *}$ are not ordered, but if they were, then the optimal classes would more and more spread over the space, which would provide nice progressive refinement of the optimal cuts. To achieve this goal, we need the supplementary axiom of *scale increasingness*(13), which yields the following axiomatic:

Definition Let $\Pi(E)$ stand for the set of all cuts of the hierarchy $H(E)$, and $\Pi(S)$ for all cuts of the sub-hierarchy $H(S)$. We call *climbing family of energies* any family $\{\omega^\lambda, \lambda \in \Lambda\}$ of energies over $\Pi(E)$ which satisfies the three following axioms, valid for $\omega^\lambda, \lambda \in \Lambda$ and for all $\pi \in \Pi(S)$, $S \in \mathcal{S}$:

i) each ω^j is *h-increasing,*

ii) each ω^j is *singular,*

iii) the $\{\omega^\lambda\}$ are *scale increasing,* i.e. for $\lambda \leq \mu$, each support $S \in \mathcal{S}$ and each partition $\pi \in \Pi(S)$, we have that

$$\lambda \leq \mu \text{ and } \omega^\lambda(S) \leq \omega^\lambda(\pi) \Rightarrow \omega^\mu(S) \leq \omega^\mu(\pi), \qquad \pi \in \Pi(S), \quad S \in \mathcal{S}. \quad (13)$$

Axiom *i)* compares the same energy at two different levels, whereas axiom *iii)* compares two different energies at the same level. The relation (13) shows that as λ increases, the ω^λ preserves the ordering of the energies between the nodes of hierarchy H and their partial partitions. In particular, if ω_0 is h-increasing and singular, and if $\{\omega^\lambda, \lambda \in \Lambda\}$ is a climbing family, then the two families $\{\lambda\omega_0, \lambda \in \Lambda\}$ and $\{\omega^\lambda + \omega_0, \lambda \in \Lambda\}$ are climbing.

5.2. Ordering of the optimal cuts

The climbing energies satisfy the very nice property of refinement of optimal cuts with respect to the parameter λ:

Theorem 5.1. *Let $\{\omega^\lambda, \lambda \in \Lambda\}$ be a family of climbing energies, and let $\pi^{\lambda *}$ (resp. $\pi^{\mu *}$) denote the optimal cut of hierarchy H according to the energy ω^λ (resp. ω^μ). Then the family $\{\pi^{\lambda *}, \lambda \in \Lambda\}$ of the optimal cuts generates a hierarchy H^* of partitions, i.e.*

$$\lambda \leq \mu \quad \Rightarrow \quad \pi^{\lambda *} \leq \pi^{\mu *}, \qquad \lambda, \mu \in \Lambda. \quad (14)$$

Proof. Given λ, there exists a unique optimal cut $\pi^{\lambda *}$ of hierarchy H (axiom *ii*)). At the class S of $\pi^{\lambda *}$ that contains point x we have $\omega^\lambda(S) \leq \omega^\lambda(\pi)$ for every partial partition π of S. Then the scale increasingness (13) implies that $\omega^\mu(S) \leq \omega^\mu(\pi), \mu \geq \lambda$. By h-increasingness of ω^μ, and by uniqueness, class S is thus temporary optimal for ω^μ, which means that the class of the final optimal cut for ω^μ at point x covers S. □

Family $\{\omega^\lambda, \lambda \in \Lambda\}$ is climbing in two senses: for each λ the energy climbs pyramid H up to its optimal cut (h-increasingness), and as λ varies, it generates a new pyramid to be climbed (scale-increasingness). Relation (14) expresses "causality" as described by Koenderink [19], i.e. the structure found at scale μ should find a cause at scale λ for $\lambda \leq \mu$.

6. Algorithms for optimal cuts

This section is devoted to three algorithms. First, to find the optimal cut for a given climbing energy ω, while the second, is to determine the scales of appearance of the classes of the provisional optimal cuts, and finally the third to find the family of optimal cuts for a climbing family of energies.

Figure 4: Optimal Cuts Pyramids: Optimal cuts using energy in equation (20), shown for different λs. Original image 25098, $\lambda_{25098} = 0(\text{leaves}), 3000, 8000$ and Original image 169012, $\lambda_{169012} = 0(\text{leaves}), 400, 10000$.

The hierarchy is organized as indexed set of classes present in each level (or partition). For each class S we introduce 2 operations. Firstly we can access children of S by the $ChildOf(S)$ operation, and secondly, we can access the parent S by $ParentOf(S)$ operation. Computationally, the h-increasingness condition (5) allows us to reach the optimal cut in one ascending pass (linear in the number of classes in H), by generalizing the algorithm by Guigues [17] to all h-increasing energies. The law of composition is referred to by $ComposeFunc$ procedure and can be the sum, infimum or supremum with their scalar weighted versions as explained earlier in propositions 4.5 and 4.6. The $ComposeFunc$ is used to calculate the composition of two energies as exemplified in Eq. (15), where the two energies $\omega_\phi(S)$ and $\omega_\partial(S)$ that represent fidelity term and a regularization term, respectively. These can be any pair of energies as further demonstrated in section 7.

The two types of $ComposeFunc$: $ComposeFunc(\omega_\phi(S),\ \omega_\partial(S), \lambda)$ or $ComposeFunc(\omega(S))$

- Addition: $\displaystyle\sum_{S_i \in childof(S)} \omega_\phi(S_i) + \lambda\omega_\partial(S_i)$

- Supremum: $\bigvee_{S_i \in childof(S)} \omega(S_i)$
- Infimum: $\bigwedge_{S_i \in childof(S)} \omega(S_i)$

6.1. Lambda list algorithm

The lambda list algorithm developed here is the same as in [17]. The scale of appearance of each class S based on the scale parameter λ is described here. The purpose of the algorithm is to extract all the scale parameter λ as described in equation 15 values at which a particular class(node) of the hierarchy appears. This is also the value of λ at which the child classes will disappear. Thus this algorithm depends on the energies $\omega_\partial(S)$ and $\omega_\varphi(S)$ for each class, and the composition function $ComposeFunc$. The elementary step of the algorithm as described in the algorithm 2, is to compare the parent energy $\omega(P) = \omega_\partial(P) + \lambda\omega_\varphi(P)$ w.r.t the composition of energies of the children $\omega(C) = ComposeFunc(\omega(c_i))$, where c_i refers to of each child whose energy is $\omega(c_i) = \omega_\partial(c_i) + \lambda\omega_\varphi(c_i)$. The composition shown in toy example in figure 5 corresponds to addition.

The appearance of scales of each class with respect to its children, here is calculated by a simple linear inequality, $\Lambda(P) = \omega(P) \leq \omega(C)$, which gives the valid value λ and thus the list of all possible optimal cuts. This result is formally described in 5.1. The $LambdaList$ algorithm implemented in this paper uses a simple quantization of the increment $\Delta\lambda$, while to obtain all values without prior knowledge requires a global branch bound search for the optimal λs.

A set of the optimal cuts belonging to the optimal cut pyramid is shown for two images in figure 4. The leaves (finest partition) is obtained at $\lambda = \lambda_{min}$, at which all the leaves classes are present. Partitions for higher values of λ are shown, which shows the series of partitions in the optimal cut pyramid. The optimal cut algorithm can be seen as a max-flow problem on directed planar graphs which are trees (see section 9). While for a directed planar graph the complexity is $nlog(n)$, the complexity for a tree 1 is at worst linear. The complexity of the algorithm 3 is linear in the number of λs of the input pyramid [17].

Algorithm 1: Optimal Cut algorithm: $OptimalCut(H, \lambda, \omega^{\lambda}(S))$

Data: H, Scale parameter λ, Energies $\omega_{\phi}(S)$, $\omega_{\partial}(S)$ for each class S.
Result: Optimal cut π^*
begin
  $NumLevels \longleftarrow |s|$ Assign the maximum number of levels (partitions) in the hierarchy H. This integer index is given by the number of unique values of saliency function s
  $NumClass \longleftarrow 0$ Initialized to 0. Assign the maximum number of classes in a given partition (level) of the hierarchy H
  $\pi^* \longleftarrow \emptyset$ Initialise the optimal cut to empty set
  for $level \in [2, NumLevels]$ **do**
    $NumClass(level) \longleftarrow |\pi(level)|$
    for $S \in [1, NumClass(level)]$ **do**
      $ChildList \longleftarrow ChildOf(S)$
      $UpperParent \longleftarrow ParentOf(S)$ Parent of Current class
      $\Omega_{Class} \longleftarrow \omega_{\phi}(S) + \lambda\omega_{\partial}(S)$
      $\Omega_{Child} \longleftarrow ComposeFunc(\omega_{\phi}^{\lambda}(S_{Child}), \omega_{\partial}^{\lambda}(S_{Child}), \lambda)$ Where $S_{Child} \in ChildList$.
      if $(\Omega_{Class} \leq \Omega_{Child})$ **then**
        $ChildOf(S) \longleftarrow \emptyset$
        else
          $ChildOf(UpperParent) \longleftarrow$ $ChildOf(UpperParent) \bigcup S_{child} \setminus S$ where $S_{child} \in ChildList$
        end
      end
    end
  end
  $\pi^* \longleftarrow \bigcup S$ Where $S \in H$ represents classes that are not deleted.
end

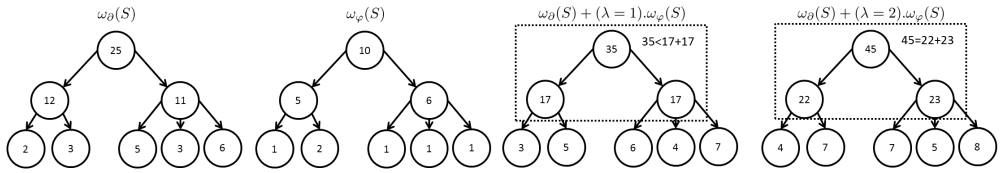

Figure 5: A toy example of the lambda list algorithm is shown. Two energies ω_∂ and ω_φ as described in equation 15, which are represented by toy values in the first two trees. The final energies $\omega_\partial(S) + \lambda\omega_\varphi(S)$ for two trees are shown for $\lambda = 1$ and $\lambda = 2$. For this example, we consider the composition of child energies by addition, $ComposeFunc(\omega(c_i)) = \sum_i(\omega(c_i)$. The scale of appearance of the root node, shown in dashed lines, is shown to be $\Lambda(root) = 2$, which is value of λ for which the parent energy equals the child energy, rendering the parent class (in this case the root) optimal.

Algorithm 2: Lambda List: $LambdaList(H, \omega_\phi(S), \omega_\partial(S))$

Data: H, Energies $\omega_\phi(S)$, $\omega_\partial(S)$ for each class S.
Result: Lambda List Λ of values which correspond to λs at which parent class S appears.

```
begin
    NumLevels ⟵ |s| Assign the maximum number of levels
    (partitions) in the hierarchy H. This integer index is given by the
    number of unique values of saliency function s
    NumClass ⟵ 0 Initialized to 0. Assign the maximum number of
    classes in a given partition (level) of the hierarchy H
    λ ⟵ 0
    while All classes have not appeared do
        for level ∈ [2, NumLevels] do
            NumClass(level) ⟵ |π(level)|
            for S ∈ [1, NumClass(level)] do
                ChildList ⟵ ChildOf(S)
                Ω^λ_{child−φ}(T) ⟵ 0
                Ω^λ_{child−∂}(T) ⟵ 0
                for T ∈ ChildList do
                    Ω^λ_{child−φ}(T) ⟵ Ω^λ_{child−φ}(T) + ω_φ(T)
                    Ω^λ_{child−∂}(T) ⟵ Ω^λ_{child−φ}(T) + ω_∂(T)
                end
                Update the Lambda List. This is the value of λ at which
                parent S is optimal(equal in energy) w.r.t to its children
                Λ(S) ⟵ (Ω^λ_{child−φ}(T) − ω_φ(S)) / (ω_∂(S) − Ω^λ_{child−∂}(T))
            end
        end
    end
end
```

Algorithm 3: Optimal Cut Pyramids

Data: Input hierarchy of partitions H, represented by saliency s,
Energy $\omega^\lambda(S)$ parametrized by λ for each class S.

Result: Optimal cut pyramid H^*

```
begin
    H* ⟵ ∅
    j ⟵ 1, i ⟵ 1 indexing starts at 1
    Λ ⟵ LambdaList(H)
    π^i ⟵ OptimaCut(H, Λ(i), ω^{Λ(i)})
    while |π^i| > 1 do
        i ⟵ i + 1
        π^i ⟵ OptimaCut(H, Λ(i), ω^{Λ(i)})
        if π^i ∉ H* then
            H*(j) ⟵ π^i
            j ⟵ j + 1
        end
    end
end
```

Algorithm 3 is not an efficient implementation, we present it for the sake of pedagogy to demonstrate how scale increasingness works. One has a complexity which is linear according to L.Guigues implementation.

7. A few useful h-increasing energies and experiments

Following the two modes of composition given by proposition 4.5, we now review two families of climbing energies obtained by addition and by supremum.

7.1. Additive energies

Scale increasingness. Proposition 4.5 has shown that ω is h-increasing. Concerning the scale increasingness, a supplementary assumption of affinity is needed [16]. It is obtained by decomposing $\omega^\lambda(S)$ into a linear function of λ:

$$\omega^\lambda(S) = \omega_\varphi(S) + \lambda\omega_\partial(S) \qquad S \in \mathcal{S}. \tag{15}$$

When the energy ω_∂ is sub-additive, i.e.

$$\omega_\partial(\bigcup_{1\leq u\leq q} T_u) \leq \sum_{1\leq u\leq q} \omega_\partial(T_u), \tag{16}$$

then the family is obviously scale increasing, since

$$\omega_\partial(S) = \omega_\partial(\bigcup_{1\leq u\leq q} T_u) \leq \sum_{1\leq u\leq q} \omega_\partial(T_u) = \omega_\partial(\pi(S)).$$

Conversely, L. Guigues showed that the condition (16) is necessary for scale increasingness [16] [17]. Below we summarise three different examples of additive energies.

7.2. Three examples

Mumford and Shah energy. It is the most popular additive energy, and historically the first [25]. One can find an exhaustive study of this functional in [26]. We write it for the Euclidean plane, and suppose that the edges are rectifiable simple arcs. Its first term, called *fidelity term* (ω_φ in (15)), sums up the quadratic differences between f and its average $m(T_u)$ in the various T_u, and the second term, called *regularization term*, weights by λ^j the lengths ∂T^i of the frontiers of all T_u, i.e.

$$\omega^j(\pi(S)) = \sum_{1\leq u\leq q} \int_{x\in T^u} \parallel f(x) - m(T_u) \parallel^2 + \lambda^j \sum_{1\leq u\leq q} (\partial T_u) \tag{17}$$

where the weight λ^j is a numerical increasing function of the level number j. Both increasingness relations (5) and (13) are satisfied by the family of energies Rel.(17), which therefore are climbing. Here the term ω_∂ involves the arc length function, but it is not the only choice. One can also think about another $\omega_\partial(S)$, which reflects the convexity of A.

Additive energy by convexity. Consider in $\mathbb{R}^2$ a connected set S without holes and with a non zero curvature everywhere on ∂S. Let $d\alpha$ be the elementary rotation of its outward normal along the element du of the frontier ∂S. As the curvature $c(u)$ equals $d\alpha/du$, and as the total rotation of the normal around ∂S equals 2π, we have

$$2\pi = \int_{c\geq 0} c(u)du - \int_{c<0} |c(u)|\, du.$$

When dealing with partitions, the distinction between outward and inward vanishes, but the parameter

$$\nu(X) = \frac{1}{2\pi} \int_{\partial S} |c(u)|\, du \tag{18}$$

still makes sense. It reaches the minimum value 1 when set S is convex, and increases with the degree of concavity. For a starfish with five pseudo-podes, it values around 5. Now $\nu(S)$ is sub-additive for the open parts of contours, therefore it can participate as a regularity term in an additive energy. In digital implementation, the angles between contour arcs must be treated separately (since sub-additivity applies on the open parts).

Additive energies by active contours. The active contours aim to match regular closed curves with the zones of maximum variation in an image [18] [11]. The energies we view are particular cases of active contours adapted to hierarchies, and derive from the approach proposed by Y. Xu at Al. [40]. The main idea is the following: each node $S \in H$ is dilated and eroded by a disc B, and the two crowns $S \oplus B \backslash S$, and $S \backslash\ S \ominus B$ are compared. This comparison stands for the fidelity term in Rel. (15), and a function of the curvature (e.g. Rel. (18)) stands for the regularity term. One goes from sets to partial partitions by additivity, according to the relation (11).

The simplest comparative energy is given by the difference of a given energetic function f on the two crowns:

$$\omega_\varphi(S) = \mid \textstyle\int_{(S \oplus B \backslash S)} f(x)dx - \int_{(S \backslash S \ominus B)} f(x)dx \mid, \qquad S \in \mathcal{P}(E). \tag{19}$$

It can be expressed in a dimensionless form by putting:

$$\widetilde{\omega}_\varphi(S) = \mid \frac{\int_{(S \oplus B \backslash S)} f(x)dx - \int_{(S \backslash S \ominus B)} f(x)dx}{a(S)} \mid, \qquad S \in \mathcal{P}(E),$$

where $a(S)$ denotes the area of S. When the absolute value bars are removed, the both energies ω_φ and $\widetilde{\omega}_\varphi$ become sub-additive. Alternatively, the energy ω_φ proposed in [40] is the sum of the variances of f in the two crowns, divided by the variance of f in the union of these two crowns.

For the regularity term ω_∂ of the energy (15), one classically takes the above function ν of Rel. (18), which is scale increasing and generates the climbing family $\{\omega_\varphi + \lambda\omega_\partial\}$.

7.3. Demonstration

Optimal cut example for color image segmentation. We aim to find an optimal simplified version of a colour image f, constrained by compression rate. A hierarchy H has been obtained on input image figure 2, by segmentations of the scalar luminance $l = (r+g+b)/3$ based on flooded watersheds [14]. In each class S of H, the simplification consists in replacing the function f by its colour mean, i.e. the means of the three channels over S. Note that this colour mean does not intervene directly in the three energies (20) to (22), but rather in the display of the optimal cut. We use the energy $\omega(S)$, as defined in equation (15) to demonstrate different optimal cuts.

For the first experiment, as the data fidelity term $\omega_\varphi(S)$ we take the variance of the luminance for each class S of hierarchy H (first term of equation (20)). The regularization term $\omega_\partial(\pi)$, is equal to the contour length $\mid \partial S \mid$, plus 24 bits for the average color of S. This gives the energy $\omega_{lum}(S)$, whose result is shown in figure 6 (left).

$$\omega_{lum}(S) = \sum_{x \in S} \parallel l(x) - \bar{l}(S) \parallel^2 + \lambda(24 + \mid \partial S \mid), \tag{20}$$

In a second experiment depicted in figure 6(right), we separate each colour vector (r, g, b) into two components: the vector luminance $\overrightarrow{l} = (l, l, l)$ which gives the gray scale, plus the orthogonal chrominance vector $\overrightarrow{c} = (r-m, g-m, b-m) = (c_1, c_2, c_3)$ whose module is the saturation.The fidelity term of the energy is now the sum of the variances of the components of $\overrightarrow{c}$ over S as shown in (21).

$$\omega_{chrom}(S) = \sum_{x \in S, 1 \leq i \leq 3} \parallel c_i(x) - \bar{c}_i(S) \parallel^2 + \lambda(24 + \mid \partial S \mid), \tag{21}$$

The principle idea here is to show that the parameters of the image involved in ω (namely the chrominance) can be completely de-correlated from those used for generating the hierarchy (namely the luminance). We observe in figure 6 (right) that the reeds in front of the female duck are now correctly segmented, and that the water in the background has lesser detail.

$$\omega_{Texture}(S) = \omega_{chrom}(S) + \sum_{S\prime \in siblings(S)} \frac{\mu}{\sigma^2(Area(S\prime))}, \tag{22}$$

This leads to a third experiment, depicted in figure 7, based on energy in equation (22). This experiment is to extract regular textures. The first term

Figure 6: Optimal cuts by using variance of luminance(left), chrominance(right)

Figure 7: Optimal Cuts for texture using variance of chrominance for scale $\lambda = 100$: Left, input Image, middle and right, cuts for parameters at $\mu = 10^{12}$ (low uniformity) and $\mu = 10^{14}$ (high uniformity), in Eq. 22.

of (22) is the same as (21). The second term in (22) is inversely proportional to the uniformity of the sizes of the children. It is measured by the variance of their areas. This is done also with the constraint that the variance of the chrominance vector is reduced over the partitions of pyramid produced from the luminance vector l. Intuitively, texture features are formulated into this multi-scale framework where the optimal scale parameter combines the effect of chrominance and structure of texture into one global energy function, thus showing the flexibility of the framework.

We conclude this section by a nice interpretation, already pointed out in Ph. Salembier and L. Garrido [33], and L. Guigues [17]. According to Lagrange formalism, one classically reaches an optimum under constraint $\omega(S)$ by means of a system of partial derivatives. Now remarkably our approach, in both cases, replaces the computation of derivatives by *climbings*.

Suppose that the value of the regularization term $\omega_\partial(\pi)$ is imposed as a fixed cost K. The cost function $\omega_\partial(\pi)$ in the optimal cut decreases as λ^j increases, since there are lesser contours. We can thus climb the pyramid of the optimal cuts and stop when $\omega_\partial(\pi)$ is the closest to K.

7.4. Sup-generated energies

The composition by supremum appears in several circumstances. For example, when dealing with the variation of a numerical function over a partial partition versus that of its classes, or in the problems of proximity to a ground truth, where the farthest distance form a point of some partial partition π to a ground truth set is the supremum of the farthest distances for the classes of π [21]. Proposition 4.5 has shown that ω is h-increasing.

Binary energies composed by supremum. The simplest $\vee$-energies are indeed the binary ones, which take values 1 and 0 only. Consider a binary $\vee$-energy ω such that for all $\pi, \pi_0, \pi_1, \pi_2 \in \mathcal{D}(E)$ we have

$$\begin{aligned} \omega(\pi) = 1 \quad &\Rightarrow \quad \omega(\pi \sqcup \pi_0) = 1, \\ \omega(\pi_1) = \omega(\pi_2) = 0 \quad &\Rightarrow \quad \omega(\pi_1 \sqcup \pi_0) = \omega(\pi_2 \sqcup \pi_0). \end{aligned}$$

This binary $\vee$-energy is obviously h-increasing. The Soille-Grazzini minimization provides an example of this type [38] [39]. A numerical function f is now associated with hierarchy H. Consider the range of variation $\delta(S) = \max\{f(x), x \in S\} - \min\{f(x), x \in S\}$ of f inside set S, and the h-increasing binary energy $\omega^k(\langle S \rangle) = 0$ when $\delta(S) \leq k$, and $\omega^k(\langle S \rangle) = 1$ when not. Compose ω according the law of the supremum, i.e. $\pi = \sqcup \langle S_i \rangle \Rightarrow \omega^k(\pi) = \bigvee_i \omega^k(\langle S_i \rangle)$. Then the class of the optimal cut at point $x \in E$ is the larger class of H whose range of variation is $\leq j$. When the energy ω^k of a father equals that of its sons, one keeps the father when $\omega^k = 0$, and the sons when not. As k varies a climbing family is generated.

Ordered energy composed by supremum. Here is an example of ordered energy due to H.G.Akcay and S. Aksoy [1] who study airborne multi-bands images and introduce (up to a small change) $\mu(S)$ =Area $(S) \times$ (mean of all standard deviations of all bands in S). They work with energy maximization. Allocate a non negative measure $\mu(S)$ to each node of a hierarchy H, where μ takes its values in a partially ordered set M, such as a color space. The energy ω is ordered by the two conditions

$$\omega(S) \leq \omega(S') \;\Leftrightarrow\; S \supseteq S' \;\&\; \mu(S) \geq \mu(S') \quad S\ , S' \in \mathcal{P}(E),\ \mu \in M. \tag{23}$$

The node S^* of the optimal cut at point x is the highest more energetic than all its descendants. The optimal cut π^* is obtained in one pass, by Guigues' algorithm [17].

Figure 8 shows optimal cuts for three different laws of composition. In a) the additive mode chooses the father S, when $\omega(S) \leq \sum \omega(T_j)$. In b) the mode by supremum chooses the S, when $\omega(S) \leq \vee\, \omega(T_j)$. Finally, in c) one takes the largest node which is more energetic than all its descendants (maximization of ω).

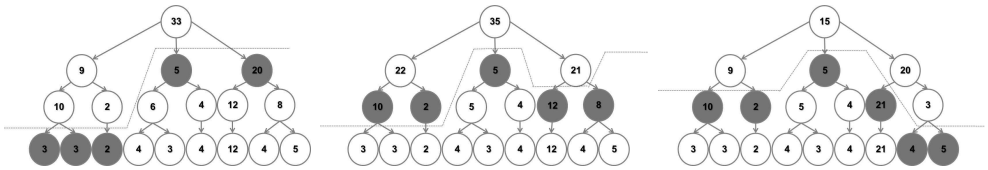

Figure 8: Optimal cuts for composition laws: addition, supremum and refinement. According to the application other laws may be used e.g. both supremum and infimum for the proximity of ground truth with Haussdorff distances [21].

Composition of $\vee$*-generated energies.* Though the weighted supremum of $\vee$-generated energies is h-increasing (Prop. 4.5), the infimum is not. In practice, this half-result is nevertheless useful, since the $\vee$, paradoxically, expresses the intersection of criteria. For example, when the function f to optimize is colour, one can take for energies:

- $\omega_1(S) = 0$ when range of luminance in $S < k_1$, and $\omega_1(S) = 1$ when not,
- $\omega_2(S) = 0$ when range of saturation in $S < k_2$, and $\omega_2(S) = 1$ when not.

Then the h-increasing energy $\omega_1(S) \vee \omega_2(S) = 0$ when S is constant enough for both luminance and saturation.

8. Partial optimizations

Covering the whole space with some optimal partition is not always an aim. Some studies require doing it, but in others ones the regions of interest are limited, and clearly marked out by the context. Moreover, the leaves partition often includes a good many classes due to noise. And thirdly, the hierarchies generated by connected filters may comprise a large number of singleton classes [32]. For example, Figure 9 b) and c) depict the flat zones obtained by an alternating filter by reconstruction acting on the pepper image a). All black pixels indicate the singleton flat zones. When climbing the hierarchy, most of these point classes are covered by extended classes which are more significant.

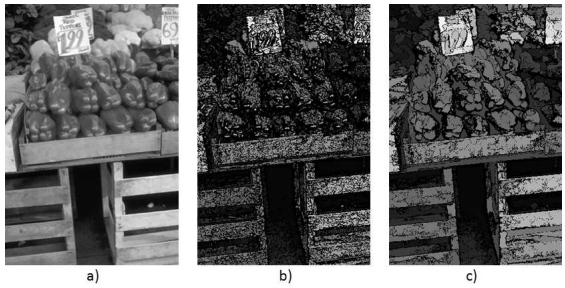

Figure 9: Alternated sequential filtering of sizes 1 and 5 of image 25098 from Berkeley Database

In other situations, some classes may be considered as non relevant because they are too small, or too large, or too far from the zone of interest, or of a non wanted hue, etc... In all cases, they are clearly identified, so that some label can indicate that they don't intervene when computing the optimal cut.

This subject extends the work on the theory of partial partitions [30]. Denote by $\mathcal{W}(E) \subseteq \mathcal{P}(E)$ the set of all classes which don't intervene. The energies ω must satisfy the condition that, for all families $\{S_i\} \subseteq \mathcal{P}(E)$ and all families $\{W_j\} \in \mathcal{W}(E)$ such that $(\cup_i S_i) \cap (\cup_j W_j) = \varnothing$, we have

$$\omega((\sqcup_i S_i) \sqcup (\sqcup_j W_j)) = \omega(\sqcup_i S_i).$$

The energy of the partial partition of classes $\{S_i\}$ must not change when external classes $\{W_j\}$ are added. It means that $\omega(W) = 0$ when the law of composition invoved in ω is the sum or the supremum, and that $\omega(W) = \infty$ when it is the infimum. When ω is h-increasing, the computation of the optimal cut is unchanged, but now results in a partition which may contain W classes.

9. Flow or Optimal Cut on Hierarchy?

It is now instructive to go back to the alternative approaches by min-cut max-flow [7], or by conditional random fields (CRFs), that we quoted in the introduction. One can notice that:

1. The CRFs and min-cut max-flow formulation represent spatial interaction between pixels which is restricted to a unitary neighborhood, and the increasing complexity may not be always advantageous [22].

2. Hierarchical methods provide a lower combinatorial complexity while supplying intuitive segmentations. In addition, the construction of a hierarchy and of the ulterior energy ω may use independent pieces of information (e.g. in section 5, luminance based energy, versus chrominance and texture).

Here we compare the two methods of optimization we contrasted in the introduction. Is it possible to interpret the structure of the data in a hierarchy H, and the notions involved in the search for its optimal cut, in terms of maximum flow problem, or dually of optimal cut, in a graph G?

The definition of a flow through G requires the data of a source and a sink. The particular shape of a pyramid leads us to take for source the family A of all leaves, and for sink the whole space E. In flows, capacities are often allocated to the edges, and sometimes to the vertices. For the sake of comparison, in case of a hierarchy H, we will take the nodes. Now, in the graph case, one wants to maximize the flow, whereas above, both additive and supgenerated energies were the matter of minimizations. We must choose, and from now on we decide to maximize the hierarchical energies, i.e. to invert the ordering relations (e.g. in comparisons father/sons of the $h-$increasing case).

In a pyramid, each leaf a is connected to the root E by a unique path $[a, .., E]$, strictly increasing, and different for each leave. For example, in Figure 10 we demonstrate a toy example with sample energies shown on a dendogram. Each node is given a capacity, which appears within it, as shown in figure. As long as two paths in this graph(tree) have no common node, the flows they carry are independent, and upper bounded by the lowest capacity along the portion where they are disjoint. When two such lines meet at some node, e.g. the two paths $[a_1, .., S]$ and $[a_1, .., S]$ which meet in S in then one must adopt some law for composing them, which is exactly what the optimal cut algorithm performs in the dynamic program.

Consider for example the additive energies, which are the most similar to the flows over a directed graph. In this additive case, the capacities $\omega(T_i\}$ of the sons $\{T_i\}$ of S are added, and compared to $\omega(S)$. The min of $\sum \omega(T_i\}$ and $\omega(S)$ gives the provisional capacity of the flow in S, and one pursues the climbing. At the end, the nodes of the optimal cut are those which are

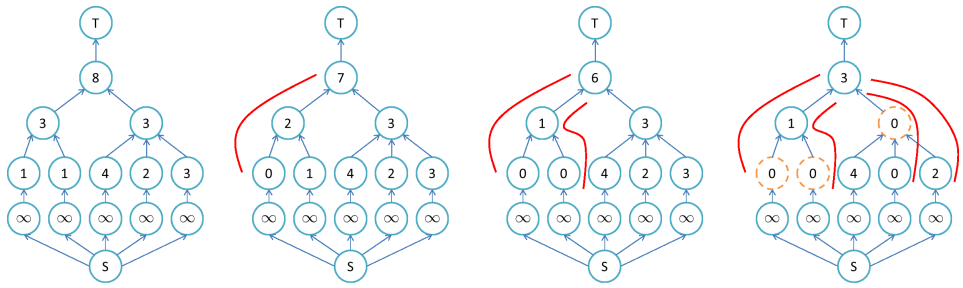

Figure 10: Flow on Hierarchy: The digram shows the min-cut for a planar graph(a tree) representing the hierarchy. The source S is connected to all the leaves by an infinite weight to force the flow through all leaves, while the sink is fixed at the root of the hierarchy. The iterations of the augmenting flow method are shown, where the minimum value on each path is subtracted from each node in the path, up till the point where we obtain a cut that separates S and T. Each augmenting flow step saturates necessarily one edge of the tree. The max-flow optimization is equivalent to the climbing optimization.

labelled 0, as depicted in Figure 10. Finally, we exactly obtain a min-cut in the graph-cut sense, but presented in another formalism, and we can state:

Proposition 9.1. *Given an additive energy ω the optimal cut of a hierarchy H is exactly the min-cut for sources located at all leaves and a sink located at the root.*

10. Conclusion

We have introduced a new method called hierarchical cuts, for finding optimal segmentations in hierarchies, which turns out to be both global and local. The approach rests on the energetic lattice of the cuts of a hierarchy (proposition 3.2), which is obtained for climbing energies, i.e. h-increasing and singular. This lattice structure results in unique optimal cuts characterized by the proposition 4.3. The fast algorithms 1 to 3 also derive from the same basic concepts.

It appeared that these two ideas were a common basis for several partition optimization methods (those which precisely work well), developed by [25], [33], [17], [38], [1], [3], [6], [10], [20], among others.

Then we introduced the climbing families of energies which provide ordered optimal cuts, thus giving a scale space interpretation to the approach.

More results at http://www.esiee.fr/ kiranr/HierarchOpt.html

Finally we demonstrated how to formulate multiple constraint functions over the image space and obtain different optimal segmentations. Two examples with colour image segmentation and texture enhancement were shown.

In future applications we aim to look at problems of evaluation of segmentation from hierarchies, labelling on hierarchies and also explicit the links to clustering. In addition we are currently developing some aspects of wavelet decompositions in the said climbing framework.

Acknowledgements. We are thankful to C. Ronse for his critical reading of the manuscript, and to F. Meyer for his helpful comments on the converse of theorem 4.1, and to B. Perret for his crucial remarks on theorem 3.1.

References

[1] Akcay, H.G., Aksoy, S.: Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE T. Geoscience and Remote Sensing* 46 (2008) pp.2097-2111

[2] Angulo J., Serra, J., Modeling and segmentation of colour images in polar representations *Image and Vision Computing* 25 (2007) 475-495.

[3] Arbelaez, P. *Une approche métrique pour la segmentation d'images*, Phd thesis, Univ.of Paris Dauphine, Nov. 2005

[4] Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour Detection and Hierarchical Image Segmentation. *IEEE PAMI 33* (2011)

[5] Bagon S., Boiman, O., Irani M., What is a Good Image Segment? A Unified Approach to Segment Extraction ECCV (2008)

[6] Ballester, C., Caselles, V., Igual, L., Garrido, L.: Level lines selection with variational models for segmentation and encoding. *JMIV 27*, 5–27 (2007)

[7] Boykov Y., Veksler O., Zabih R.: Fast Approximate Energy Minimization via Graph Cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23(11): 1222-1239 (2001)

[8] Benzécri J. P., *L'Analyse des Données. Tome I : La Taxinomie.* Dunod, Paris, 4eme edition, 1984.

[9] Bell, E. T. "Exponential Numbers." Amer. Math. Monthly 41,

[10] Cardelino, J., Caselles, V., Bertalmío, M., Randall, G.: A contrario hierarchical image segmentation. In: *IEEE ICIP 2009*, Cairo, Egypt (2009)

[11] Chan T.F. and Vese L.A., Active contours without edges, *IEEE Trans. on Image Proc.*, vol. 10, no. 2, pp. 266–277, 2001.

[12] Cousty J., Najman L., Serra J. Raising in Watershed Lattices, *2008 IEEE Intern. Conf. on Image Processing, ICIP 2008*, San Diego 13 -17 Oct. 2008.

[13] Diday E., Lemaire J., Pouget J., Testu F., *Eléments d'analyse des données* .Dunod, Paris, 1982

[14] Cousty, J., Najman, L.: Incremental Algorithm for Hierarchical Minimum Spanning Forests and Saliency of Watershed Cuts. *ISMM 2011. LNCS, vol. 6671* Springer, Heidelberg (2011) 272–283

[15] Gorelick L., Schmidt F. R., Boykov Y., Delong A., Ward A., Segmentation with non-linear regional constraints via line-search cuts, ECCV2012, LNCS 7572.

[16] Guigues L., *Modéles multi-échelles pour la segmentation d'images.*Thése doctorale Université de Cergy-Pontoise, décembre 2003.

[17] Guigues L., Cocquerez J.P., Le Men H., Scale-Sets Image Analysis, *Int. Journal of Computer Vision* 68(3), 289-317, 2006.

[18] M. Kass, A. Witkin, and D. Terzopoulos, Snakes : Active contour models,*IJCV*, vol. 1, pp. 321–331, 1987.

[19] Koenderink J.J., The structure of images Biological cybernetics 50 (5), 363-370

[20] Koepfler, G. Lopez, C. Morel, J.M., A multiscale algorithm for image segmentation by variational method,SIAM journal on numerical analysis,31,1,282-299,1994,SIAM

[21] Kiran B.R., Serra J., Ground truth energies for hierarchies of segmentations, Accepted ISMM 2013.

[22] Lucchi, A., Li, Y., Bosch, X.B., Smith, K., Fua, P.: Are spatial and global constraints really necessary for segmentation? In:*ICCV.* (2011) 9-16

[23] Meyer, F., Adjunctions on the lattice of dendrograms, submitted to *GBR2013*

[24] Meyer F., Najman L. Segmentation, Minimum Spanning Trees, and Hierarchies, in *Mathematical Morphology*, L. Najman et H. Talbot, Eds. Wiley, N.Y. 2010.

[25] Mumford D. and Shah J., Boundary Detection by Minimizing Functionals, in *Image Understanding*, S. Ulmann and W. Richards Eds, 1988.

[26] M. Morel et S. Solimini — Variational methods in image segmentation, Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser 1995

[27] Najman L., and Schmitt M., Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Trans. on PAMI,* 18(12): 1163-1173, Dec. 1996

[28] Najman L.,Talbot H., Eds**.** *Mathematical Morphology*, Wiley, N.Y. 2010.

[29] Najman, L.: On the Equivalence Between Hierarchical Segmentations and Ultrametric Watersheds. *JMIV* Volume 40(3), (2011) 231–247

[30] Ronse, C., Partial partitions, partial connections and connective segmentation. *JMIV* Volume 32, (2008) 97–125

[31] Ronse, C., Adjunctions on the lattices of partitions and of partial partitions. *AAECC*, (2010), 21(5): 343-396.

[32] Salembier P. and Serra J., Flat Zones Filtering, Connected Operators, and Filters by Reconstruction, *IEEE Transactions on Image Processing.* Aug. 1995, vol. 4, n° 8, 1153-1160 .

[33] Salembier P., Garrido L., Binary Partition Tree as an Efficient Representation for Image Processing, Segmentation, and Information Retrieval. *IEEE Trans. on Image Processing,* 2000, 9(4): 561-576.

[34] Salembier P., Connected operators based on tree pruning strategies, Chapter 7 in [28].

[35] Serra, J., A lattice approach to Image segmentation, *Journal of Mathematical Imaging and Vision*, 24: (2006) 83-130.

[36] Serra, J., Hierarchy and Optima, in *Discrete Geometry for Computer Imagery*, LNCS 6007, Springer 2011, pp 35-46

[37] J.Serra, B. R. Kiran, J.Cousty Hierarchies and climbing energies, in CIARP 2012, *LNCS 7441*, Springer-Verlag,(2012) 821–828.

[38] Soille, P., Constrained connectivity for hierarchical image partitioning and simplification. IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (2008)1132–1145

[39] Soille, P., Grazzini, J. Constrained Connectivity and Transition Regions, in *Mathematical Morphlology and its Applications to Signal and Image Processing*, M.H.F. Wilkinson and B.T.M. Roerdink, Eds, Springer 2009, pp.59-70.

[40] Xu Y., Géraud T., Najman L., Context-based energy estimator: application to object segmentation on the tree of shapes. ICIP 2012