



# Hierarchical image simplification and segmentation based on Mumford-Shah-salient level line selection

Yongchao Xu, Thierry Géraud, Laurent Najman

## ► To cite this version:

Yongchao Xu, Thierry Géraud, Laurent Najman. Hierarchical image simplification and segmentation based on Mumford-Shah-salient level line selection. 2015. hal-01287029v1

**HAL Id: hal-01287029**

**<https://hal.science/hal-01287029v1>**

Preprint submitted on 15 Mar 2016 (v1), last revised 17 May 2016 (v2)

**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.

# Hierarchical image simplification and segmentation based on Mumford-Shah-salient level line selection

Yongchao Xu<sup>\*†</sup>, Thierry Géraud<sup>\*</sup>, Laurent Najman<sup>‡</sup>

## Abstract

Hierarchies, such as the tree of shapes, are popular representations for image simplification and segmentation thanks to their multiscale structures. Selecting meaningful level lines (boundaries of shapes) yields to simplify image while preserving intact salient structures. Many image simplification and segmentation methods are driven by the optimization of an energy functional, for instance the celebrated Mumford-Shah functional. In this paper, we propose an efficient approach to hierarchical image simplification and segmentation based on the minimization of some energy functional. This method conforms to the current trend that consists in producing hierarchical results rather than a unique partition. Contrary to classical approaches which compute optimal hierarchical segmentations from an input hierarchy of segmentations, we rely on the tree of shapes, a unique and well-defined representation equivalent to the image. Simply put, we compute for each level line of the image an attribute function that characterizes its persistence under the energy minimization. Then we stack the level lines from meaningless ones to salient ones through a saliency map based on extinction values defined on the tree-based shape space. Qualitative illustrations and quantitative evaluation on Weizmann segmentation evaluation database demonstrate the state-of-the-art performance of our method.

## 1 Introduction

In natural images, meaningful contours are usually smooth and well-contrasted. Many authors (*e.g.*, [10, 6]) claim that significant contours of objects in images coincide with segments of the image level lines. The level lines are the boundaries of the connected components described by the *tree of shapes* proposed in [21], and also known as *topographic map* in [10]. Image simplification or segmentation can then be obtained by selecting meaningful level lines in that tree. This subject has been investigated in the past by [26, 6, 7]. In [20], the authors have proposed a tree simplification method for image simplification purpose based on the binary partition tree.

Classically, finding relevant contours is often tackled using an energy-based approach. It involves minimizing a two-term-based energy functional of the form  $E_{\lambda_s} = \lambda_s C + D$ , where  $C$  is the regularization term controlling the regularity of contours,  $D$  is a data fidelity term, and  $\lambda_s$  is a parameter. A popular example is the seminal work of [22]. Curve evolution methods are usually used to solve

this minimization problem. They have solid theoretical foundations, yet they are often computationally expensive.

Current trends in image simplification and segmentation are to find a multiscale representation of the image rather than a unique partition. There exist many works about hierarchical segmentations such as the geodesic saliency of watershed contours proposed in [25] and gpb-owt-ucm proposed by [2] and references therein. Some authors propose to minimize a two-term-based energy functional subordinated to a given input hierarchy of segmentations, in order to find an optimal hierarchical image segmentations in the sense of energy minimization. Examples are the works of [16, 18]. Yet, the choice or the construction of the input hierarchy of segmentations for these methods is an interesting problem in itself. [27] compared different choices of morphological hierarchies for supervised segmentation.

In this paper we propose a novel hierarchical image simplification and segmentation based on minimization of an energy functional (*e.g.*, the piecewise-constant Mumford-shah functional). The minimization is performed subordinated to the shape space given by the tree of shapes, a unique and equivalent image representation. The basis of our proposal was exposed in our previous study in [38], in which we proposed an efficient greedy algorithm computing a meaningfully locally optimal solution of the energy minimization problem. The basic idea is to take into account the meaningfulness of the set of level lines so that we very quickly get a “good” (in the sense of getting a well-simplified image while preserving the salient structures) locally optimal solution. The current paper extends this

---

<sup>\*</sup>EPITA Research and Development Laboratory (LRDE), 14-16, rue Voltaire, FR-94276 Le Kremlin-Bicêtre, France

<sup>†</sup>Department of Signal and Image Processing, Telecom ParisTech, 46 rue Barrault, 75013 Paris, France

<sup>‡</sup>Université Paris-Est, Laboratoire d’Informatique Gaspard-Monge (LIGM), A3SI, ESIEE Paris, Cité Descartes, BP 99, FR-93162 Noisy-le-Grand, France

E-mails: {yongchao.xu, thierry.geraud}@lrde.epita.fr, l.najman@esiee.fr

The paper is under consideration at Pattern Recognition Letters.

idea to hierarchical simplification and segmentation. More precisely, following the same principle but without fixing the parameter  $\lambda_s$  in the two-term-based energy, we compute an attribute function that characterizes the persistence of each shape under the energy minimization. Then we compute a saliency map, a single image presenting the complete hierarchical simplifications or segmentations. To do so, we rely on the idea of hierarchy transformation via extinction value proposed by [35] and on the framework of tree-based shape space introduced in [40]. This scheme of hierarchy transformation has been first used in [39] for a different input hierarchy and attribute function. Related algorithms were presented in [37]. The present paper extends on these ideas, focusing on the computation of an attribute function related to energy minimization.

Note that although the tree of shapes (see Section 2.1) is also a hierarchical image representation, it is very different from a hierarchy of segmentations (see Section 2.2). Each cut of the tree of shapes is a subset of the image domain, while each cut of the later one forms a partition of the image domain (see [28]). This basic difference prohibits the direct use of the classical works of finding optimal hierarchical segmentations by energy minimization to the tree of shapes. In this sense, our approach can be seen as an extension of the scale-set theory proposed in [16] to the tree of shapes.

The main contributions of this current paper are: 1) A novel hierarchical image simplification and segmentation method based on energy minimization subordinated to the tree of shapes; This extends the scale-set theory proposed by [16] to the tree of shapes. Note that the current work is quite different from our previous work in [39] in terms of input hierarchy and attribute function. In [39], the input hierarchy is a hierarchy of segmentations, and the attribute, inspired by [14], is a local measure of the interest of merging two regions. 2) A detailed discussion of the method compared with some related works. 3) Some illustrations of the proposed method on color images, as well as a quantitative evaluation of the method on a public segmentation dataset. Note that our proposal is a general framework. We have tested it with a very simple segmentation model in this paper. Despite its simplicity, we obtain results that are competitive with the ones of some state-of-the-art methods on the classical segmentation dataset from [1]. In particular, they are on par with Gpb-owt-ucm proposed in [2] on this dataset.

The rest of this paper is organized as follows: Some background information is provided in Section 2. Section 3 is dedicated to depict the proposed method, followed by some illustrations and experimental results in Section 4. Section 5 compares the proposed method with some similar works. We then conclude in Section 6.

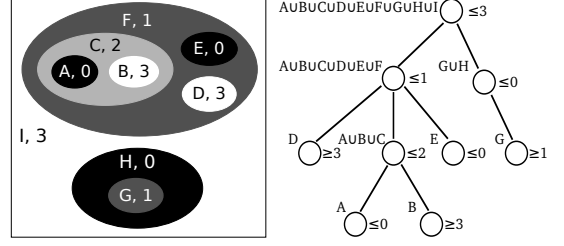


Figure 1: An image (left) and its tree of shapes (right).

## 2 Background

### 2.1 The Tree of shapes

For any  $\lambda \in \mathbb{R}$  or  $\mathbb{Z}$ , the upper level sets  $\mathcal{X}_\lambda$  and lower level sets  $\mathcal{X}^\lambda$  of an image  $f : \Omega \rightarrow \mathbb{R}$  or  $\mathbb{Z}$  are respectively defined by  $\mathcal{X}_\lambda(f) = \{p \in \Omega \mid f(p) \geq \lambda\}$  and  $\mathcal{X}^\lambda(f) = \{p \in \Omega \mid f(p) \leq \lambda\}$ . Both upper and lower level sets have a natural inclusion structure:  $\forall \lambda_1 \leq \lambda_2, \mathcal{X}_{\lambda_1} \supseteq \mathcal{X}_{\lambda_2}$  and  $\mathcal{X}^{\lambda_1} \subseteq \mathcal{X}^{\lambda_2}$ , which leads to two distinct and dual representations of an image, the Max-tree and the Min-tree.

Another tree has been introduced in [21] via the notion of shapes. A *shape* is defined as a connected component of an upper or lower level set where its holes have been filled in. Thanks to the inclusion relationship of both kinds of level sets, the set of shapes gives a unique tree, called *tree of shapes*. This tree is a self-dual, non-redundant, and complete representation of an image. It is equivalent to the input image in the sense that the image can be reconstructed from the tree. And it is invariant to affine contrast changes. Such a tree also inherently embeds a morphological scale-space (the parent of a node/shape is a larger shape). An example on a synthetic image is depicted in Fig. 1. Recently, an extension of the tree of shapes for color images has been proposed by [9] through the inclusion relationship between the shapes of its three grayscale channels.

### 2.2 Hierarchy of image segmentations or saliency maps

A hierarchy of image segmentation  $H$  is a multiscale representation that consists of a set of nesting partitions from fine to coarse:  $H = \{\mathcal{P}_i \mid 0 \leq i \leq n, \forall j, k, 0 \leq j \leq k \leq n \Rightarrow \mathcal{P}_j \subseteq \mathcal{P}_k\}$ , where  $\mathcal{P}_n$  is the partition  $\{\Omega\}$  of  $\Omega$  into a single region, and  $\mathcal{P}_0$  represents the finest partition of the image  $f$ .  $\mathcal{P}_j \subseteq \mathcal{P}_k$  implies that the partition  $\mathcal{P}_j$  is finer than  $\mathcal{P}_k$ , which means  $\forall R \in \mathcal{P}_j, \exists R' \in \mathcal{P}_k$  such that  $R \subseteq R'$ .

As a multiscale representation, a hierarchy of segmentation satisfies the most fundamental principle for multiscale analysis: the causality principle presented by [19]. From this principle, for any couple of scales  $\lambda_{s_2} > \lambda_{s_1}$ , the “structures” found at scale  $\lambda_{s_2}$  should find a “cause” at scale  $\lambda_{s_1}$ . In the case of a hierarchy of segmentation, following the work of [16], the causality principle is applied

to the edges associated to the set of partitions spanned by  $H$ : for any pair of scales  $\lambda_{s_2} > \lambda_{s_1}$ , the boundaries of partition  $\mathcal{P}_{\lambda_{s_2}}$  are in a one-to-one mapping with a subset of the boundaries of  $\mathcal{P}_{\lambda_{s_1}}$  (their “cause”). The pair  $(H, \lambda_s)$  is called an indexed hierarchy.

A useful representation of hierarchical image segmentations was originally introduced in [25] under the name of *saliency map*. A saliency map is obtained by stacking a family of hierarchical contours. This representation was then rediscovered independently by [16] through the notion of scale-set theory for visualization purposes, and it is then popularized by [2] under the name of *ultrametric contour map* for boundary extraction and comparing hierarchies. It has been proved theoretically in [23] that a hierarchy of segmentations is equivalent to a saliency map. Roughly speaking, for a given indexed hierarchy  $(H, \lambda)$ , the corresponding saliency map can be obtained by weighing each contour of the image domain with the highest value  $\lambda_s$  such that it appears in the boundaries of some partition represented by the hierarchy  $H$ . The low level (resp. upper level) of a hierarchy corresponds to weak (resp. strong) contours, and thus an over-segmentation (resp. under-segmentation) can be obtained by thresholding the saliency map with low (resp. high) value.

## 2.3 From shape-space filtering to hierarchy of segmentations

The three morphological trees reviewed in Section 2.1 and the hierarchies of segmentations reviewed in Section 2.2 have a tree structure. Each representation is composed of a set of connected components  $\mathbb{C}$ . Any two different elements  $C_i \in \mathbb{C}, C_j \in \mathbb{C}$  are either disjoint or nested:  $\forall C_i \in \mathbb{C}, C_j \in \mathbb{C}, C_i \cap C_j \neq \emptyset \Rightarrow C_i \subseteq C_j$  or  $C_j \subseteq C_i$ . This property leads to the definition of *tree-based shape space* in [40]: a graph representation  $G_{\mathbb{C}} = (\mathbb{C}, E_{\mathbb{C}})$ , where each node of the graph represents a connected component in the tree, and the edges  $E_{\mathbb{C}}$  are given by the inclusion relationship between connected components in  $\mathbb{C}$ . In [40], we have proposed to filter this shape space by applying some classical operators, notably connected operators on  $G_{\mathbb{C}}$ . We have shown that this shape-space filtering encompasses some classical connected operators, and introduces two families of novel connected operators: shape-based lower/upper leveling and shaping.

Instead of filtering the shape space, another idea is to consider each region of the shape space as a candidate region of a final partition. For example, we weigh the shape space by a quantitative meaningfulness measurement  $\mathcal{A}$ . Then each local minimum of the node-weighted shape space is considered as a candidate region of a partition. The importance of each local minimum (*i.e.*, each region) can be measured quantitatively by the extinction value  $\mathcal{E}$  proposed by [35]. Let  $\prec$  be a strict total order on the set of minima  $m_1 \prec m_2 \prec \dots$ , such that  $m_i \prec m_{i+1}$  whenever  $\mathcal{A}(m_i) < \mathcal{A}(m_{i+1})$ . Let  $CC$  be the lowest lower level connected component (defined on the shape space) that

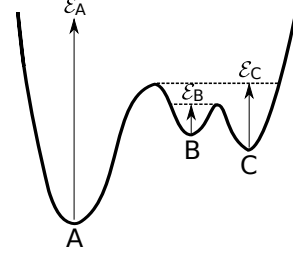


Figure 2: Illustration of the extinction values  $\mathcal{E}$  of three minima. The order is  $A \prec C \prec B$ .  $B$  merges with  $C$ ,  $C$  merges with  $A$ .

contains both  $m_{i+1}$  and a minimum  $m_j$  with  $j < (i + 1)$ . The extinction value for the minimum  $m_{i+1}$  is defined as the difference of level of  $CC$  and  $\mathcal{A}(m_{i+1})$ . An example of extinction values for three minima is depicted in Fig 2. We weigh the boundaries of the regions corresponding to the local minima with the extinction values. This yields a saliency map representing a hierarchical image simplification or segmentation. This scheme allows to transform any hierarchical representation into a hierarchical segmentation. It has been firstly used in [39], where the input hierarchy is a minimum spanning tree and the attribute is computed locally inspired from the work of [14]. For the sake of completeness, the algorithm [37] for the extinction-based hierarchy transformation is presented in Section 3.4.

## 2.4 Energy-based simplification and segmentation

There exist several works of hierarchical image segmentations based on energy minimization (see [16] and [18]). A general formulation of these methods involves minimizing a two term-based energy functional of the form  $E_{\lambda_s} = \lambda_s C + D$ .  $C$  is a regularization term,  $D$  is data fidelity term, and  $\lambda_s$  is a parameter. Let  $\{R\} = R_1 \sqcup \dots \sqcup R_n$  be a partition of the image domain. If the energy functional can be written by  $E_{\lambda_s} = \sum_{R_i \in \{R\}} (\lambda_s C(R_i) + D(R_i))$ , then  $E_{\lambda_s}$  is called an affine separable energy functional. Furthermore, if either the regularization term  $C$  decreases or the data fidelity term  $D$  increases, the energy  $E_{\lambda_s}$  is multiscale affine separable. A popular instance of such an energy functional that we will use as an example through this paper is the piecewise-constant Mumford-Shah functional proposed by [22]. For an image  $f$ , it is given by

$$E_{\lambda_s}(f, \partial\{R\}) = \iint_{\{R\}} (\tilde{f}_i - f)^2 dx dy + \lambda_s |\partial\{R\}|, \quad (1)$$

where  $\tilde{f}_i = \frac{1}{|R_i|} \iint_{R_i} f dx dy$  inside each region  $R_i \in \{R\}$ ,  $\partial\{R\}$  is the set of contour, and  $|\cdot|$  denotes the cardinality.

### 3 Hierarchical image simplification and segmentation via level line selection

#### 3.1 Main idea

The current proposal extends the method of image simplification by selecting salient level lines based on Mumford-Shah energy functional minimization, which has been presented in a preliminary version of this study in [38]. We review this non-hierarchical version in Section 3.2, using a more general multiscale affine separable energy. It has been shown in [38] that the non-hierarchical version of the proposed method belongs to the class of morphological shapings in the framework of shape-space filtering. In this current paper, we compute for each level line an attribute function related to the energy’s regularization parameter  $\lambda_s$ , and use the idea of hierarchy transformation based on extinction values and on a tree-based shape space to achieve hierarchical image simplification and segmentation. This is detailed in Section 3.3. Section 3.4 provide algorithms for the whole process.

#### 3.2 Image simplification by salient level line selection

For a given tree of shapes  $\mathcal{T}$  composed of a set of shapes  $\{\tau_i\}$ , any two successive shapes of  $\mathcal{T}$  are related by an edge reflecting the inclusion relationship, also known as the parenthood between nodes of the tree. This tree structure  $\mathcal{T}$  provides an associated partition of the image  $\{R_{\mathcal{T}}\} = R_{\tau_1} \sqcup \dots \sqcup R_{\tau_n}$ , where  $R_{\tau} = \{p \mid p \in \tau, p \notin Ch(\tau)\}$  with  $Ch(\tau)$  representing all the children of the shape  $\tau$ . We denote by  $E_{\lambda_s}(f, \mathcal{T})$  the energy functional (see Section 2.4) subordinated to the tree by considering its associated partition  $\{R_{\mathcal{T}}\}$ . This energy minimization is given by:

$$\min_{\mathcal{T}'} E_{\lambda_s}(f, \mathcal{T}'), \quad (2)$$

where  $\mathcal{T}'$  is a simplified version of  $\mathcal{T}$  by removing some shapes from  $\mathcal{T}$  and by updating the parenthood relationship.

The basic operation of the energy minimization problem given by Eq. (2) is to remove the contours of some shapes  $\{\tau\}$  included in their corresponding parents  $\{\tau_p\}$ , which triggers the update of  $R'_{\tau_p} = R_{\tau_p} \cup R_{\tau}$  for each shape  $\tau$ . The parent of its children  $\tau_{c1}, \dots, \tau_{ck}$  should also be updated to the  $\tau_p$ . Fig. 3 shows an example of a such merging operation.

Observe that the minimization problem of Eq. (2) is a combinatorial optimization. The computation of the optimum has an exponential complexity. Hence a greedy algorithm is usually applied to compute a local optimum instead of a global optimum (see also [4]). It iteratively removes the shapes to decrease the energy functional. The

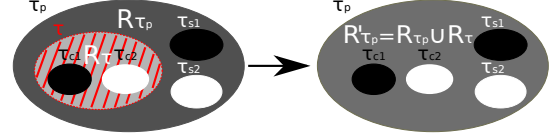


Figure 3: Suppressing the node  $\tau$  makes the “region”  $R_{\tau}$  (covered with red oblique lines) merge with  $R_{\tau_p}$ ; the result (depicted in the right image) is a simplified image.

greedy algorithm stops when no other shape can be removed that favors a decrease of the energy. The removability of a shape  $\tau$  is decided by the sign of the energy variation  $\Delta E_{\lambda_s}^{\tau}$  while  $\tau$  is suppressed. For the multiscale affine separable energy described in Section 2.4,  $\Delta E_{\lambda_s}^{\tau}$  is given by:

$$\Delta E_{\lambda_s}^{\tau} = D(R'_{\tau_p}) - D(R_{\tau}) - D(R_{\tau_p}) - \lambda_s (C(R_{\tau}) + C(R_{\tau_p}) - C(R'_{\tau_p})). \quad (3)$$

Taking the piecewise-constant Mumford-Shah functional given by Eq. (1) as an energy example, and let  $S(f, R_i)$  be the sum of value of all the pixels inside  $R_i$ , Then the functional variation  $\Delta E_{\lambda_s}^{\tau}$  is given by:

$$\Delta E_{\lambda_s}^{\tau} = \frac{S^2(f, R_{\tau})}{|R_{\tau}|} + \frac{S^2(f, R_{\tau_p})}{|R_{\tau_p}|} - \frac{S^2(f, R'_{\tau_p})}{|R'_{\tau_p}|} - \lambda_s |\partial \tau|. \quad (4)$$

If  $\Delta E_{\lambda_s}^{\tau}$  is negative, which means the suppression of  $\tau$  decreases the functional, then we remove  $\tau$ . According to Eq. (3), the removability of a shape  $\tau$  depends only on  $R_{\tau}$  and  $R_{\tau_p}$ . As the suppression of the shape  $\tau$  triggers the update of  $R_{\tau_p}$ , the removal of  $\tau$  impacts also the removability of its parent, its children and siblings. So the order of level line removal is critical. In [38], we proposed to fix the order by sorting the level lines in increasing order of a quantitative meaningfulness attribute  $\mathcal{A}$  (e.g., average of gradient’s magnitude along the level line  $\mathcal{A}_{\nabla}$ ).

Meaningful contours in natural images are usually well-contrasted and smooth. Indeed, the minimization of energy functional in Eq (1) favors the removal of level lines having small contrast (by data fidelity term) or being complex (by regularization term). So the shapes having small (resp. great) meaningfulness attribute  $\mathcal{A}$  are easier (resp. more difficult) to filter out under the energy minimization process of Eq. (2). Consequently, the level line sorting based on meaningfulness attribute  $\mathcal{A}$  provides a reasonable order to perform the level lines suppression that makes the energy functional decrease. Indeed, initially, each region  $R_{\tau}$  has only several pixels. At the beginning, many “meaningless” regions are removed, which forms more proper regions in Eq. (3) for the “meaningful” regions. The removal decisions for these “meaningful” regions based on the sign of Eq. (3) are more robust.

#### 3.3 Hierarchical salient level line selection

The parameter  $\lambda_s$ , in the multiscale affine separable energy, controls the simplification/segmentation degree for



Figure 4: Illustration of causality principle violation. Left: input image; Middle and right: randomly colorized simplified image with  $\lambda_{s_1} = 100$  and respectively  $\lambda_{s_2} = 500$ . The orange shape on top middle of the right image (surrounded by a black circle) is preserved for  $\lambda_{s_2} = 500$ , while it is removed for  $\lambda_{s_1} = 100$  in the middle image.

the method described in Section 3.2, which is however not hierarchical. Because some level line  $\tau$  may be removed with a parameter  $\lambda_{s_1}$ , but preserved for a bigger  $\lambda_{s_2} > \lambda_{s_1}$ . This contradicts the causality principle for hierarchical image simplification/segmentation described in Section 2.2. An example is given in Fig. 4. Note that the simplification algorithm of [4] is not hierarchical either.

Instead of fixing the parameter  $\lambda_s$  in the energy functional (e.g.,  $\lambda_s$  in Eq (1)), we propose to compute an individual  $\lambda_s$  for each shape of the tree following the same principle of the energy minimization. For a given  $\lambda_s$ , the removability of a shape  $\tau$  is based on the sign of energy variation  $\Delta E_{\lambda_s}^\tau$  in Eq. (3), which is a linear decreasing function w.r.t.  $\lambda_s$  (e.g., the Eq (4) for the piecewise-constant Mumford-Shah functional). When  $\lambda_s$  is bigger than some value  $\lambda_{s_{min}}$ ,  $\Delta E_{\lambda_s}^\tau$  will be negative, which implies the removal of this shape decreases the energy functional. Thus  $\lambda_{s_{min}}$  is a value of the transition for the removal decision of the underlying shape. Let us denote this value of transition as the attribute function  $\mathcal{A}_{\lambda_s}$ , which is given by:

$$\mathcal{A}_{\lambda_s}(\tau) = \frac{D(R'_{\tau_p}) - D(R_\tau) - D(R_{\tau_p})}{C(R_\tau) + C(R_{\tau_p}) - C(R'_{\tau_p})}, \quad (5)$$

For the piecewise-constant energy functional in Eq. (1), it is given by:

$$\mathcal{A}_{\lambda_s}(\tau) = \left( \frac{S^2(f, R_\tau)}{|R_\tau|} + \frac{S^2(f, R_{\tau_p})}{|R_{\tau_p}|} - \frac{S^2(f, R'_{\tau_p})}{|R'_{\tau_p}|} \right) / |\partial\tau|. \quad (6)$$

Note that for a given shape  $\tau$ , the attribute function  $\mathcal{A}_{\lambda_s}(\tau)$  defined in Eq (5) depends on  $R_\tau$  and  $R_{\tau_p}$ , which means  $\mathcal{A}_{\lambda_s}(\tau)$  is decided by the shape  $\tau$  itself, its parent, its siblings, and its children. Because the attribute function  $\mathcal{A}_{\lambda_s}$  is computed under the hypothesis that the shape  $\tau$  under scrutiny is suppressed, we also need to update  $R_{\tau_p}$ , and update the parenthood relationship for its children to its parent. These update operations will also affect the computation of  $\mathcal{A}_{\lambda_s}$  for the parent, children and siblings of  $\tau$ . So the computation order is again important. We follow the same principle as described in Section 3.2 to compute  $\mathcal{A}_{\lambda_s}$ , which is detailed as below:

*step 1:* Compute  $\mathcal{A}_{\lambda_s}$  for each shape  $\tau \in \mathcal{T}$  supposing that only the shape under scrutiny is removed, and sort

the set of shapes  $\{\tau \mid \tau \in \mathcal{T}\}$  in increasing order of shape meaningfulness indicated by an attribute  $\mathcal{A}$  (e.g.,  $\mathcal{A}_\nabla$ ).

*step 2:* Propagate the sorted shapes in increasing order, and remove the shape one by one. Compute the new value  $\mathcal{A}_{\lambda_s}$  for the underlying shape  $\tau$ , and update it if the value is greater than the older one. Update also the parenthood relationship and the corresponding information for  $R_{\tau_p}$ .

This attribute function  $\mathcal{A}_{\lambda_s}$  is related to the minimization of the energy functional. It measures the persistence of a shape to be removed under the minimization problem of Eq (2). A bigger  $\mathcal{A}_{\lambda_s}(\tau)$  means that it is more difficult to remove the shape  $\tau$ . Thus the attribute function  $\mathcal{A}_{\lambda_s}$  is also some kind of quantitative meaningfulness deduced from the energy minimization.

We use the inverse of the attribute  $\mathcal{A}_{\lambda_s}$  described above as the final attribute function:  $\mathcal{A}_{\lambda_s}^\downarrow(\tau) = \max_{\tau' \in \mathcal{T}} (\mathcal{A}_{\lambda_s}(\tau')) - \mathcal{A}_{\lambda_s}(\tau)$ . The local minima of the shape space weighted by this attribute function correspond to a set of candidate salient level lines. We make use of the scheme of hierarchy transformation described in Section 2.3 to compute a saliency map  $\mathcal{M}_\mathcal{E}$ . This saliency map  $\mathcal{M}_\mathcal{E}$  represents hierarchical result of level line selections. Each thresholding of this map  $\mathcal{M}_\mathcal{E}$  selects salient (of certain degree) level lines from which a simplified image can be reconstructed.

An example of the proposed scheme on a synthetic image is illustrated in Fig. 5. The input image in Fig. 5 (a) is both blurred and noisy. This blurring is also visible in Fig. 5 (b) that illustrates the evolution of the average of gradient's magnitude  $\mathcal{A}_\nabla$  along the contours of shapes starting from regions inside the triangle, pentagon, and square regions to the root of the tree. The evolution of the initial values of the Attribute  $\mathcal{A}_{\lambda_s}$  obtained at step 1 on the same branches of the tree are provided in Fig. 5 (c). It is not surprising that those initial values  $\mathcal{A}_{\lambda_s}$  are not effective to measure the meaningfulness of the shapes: indeed, this is due to the very small size of each region in  $\{\mathcal{R}_\mathcal{T}\}$ . The evolution of the final values of the attribute  $\mathcal{A}_{\lambda_s}$  (of step 2) is depicted in Fig. 5 (d). We can see that the significant regions are highlighted by  $\mathcal{A}_{\lambda_s}$ . This experiment also demonstrates the relevance of the increasing order of average of gradient's magnitude along the contour  $\mathcal{A}_\nabla$  as a criterion to update the value of  $\mathcal{A}_{\lambda_s}$ . The saliency map  $\mathcal{M}_\mathcal{E}$  using the attribute  $\mathcal{A}_{\lambda_s}$  and one of the possible segmentations that can be obtained by thresholding  $\mathcal{M}_\mathcal{E}$  are depicted in Fig. 5. (e) and (f).

### 3.4 Implementation

The proposed method is composed of four main steps: 1) Construction of tree of shapes; 2) Computation of attribute function  $\mathcal{A}_{\lambda_s}$ ; 3) Computation of extinction values  $\mathcal{E}$ ; 4) Computation of saliency map  $\mathcal{M}_\mathcal{E}$ .

For the tree construction, we use an immersion algorithm based on the union-find process proposed by [34]. The algorithm starts by sorting the pixels in root-to-leaf order. Then a union-find process propagates the pixels in



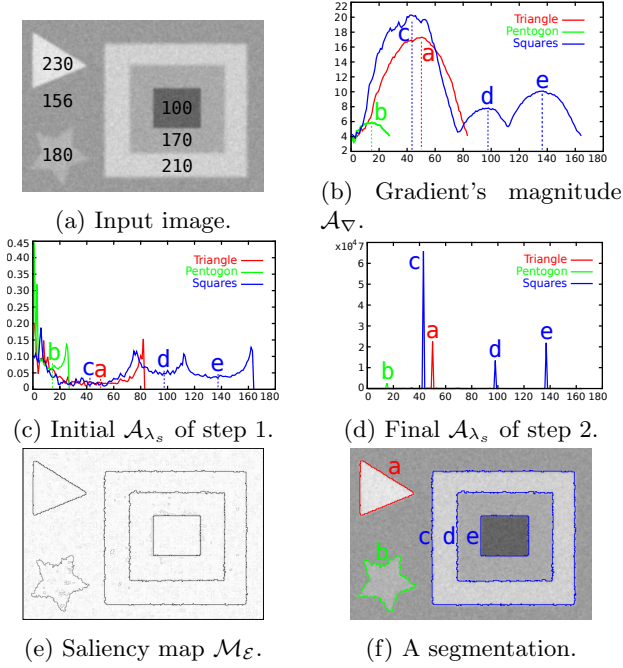


Figure 5: An example of the proposed scheme on a synthetic image. (b-d): Evolution of attribute value starting from leaf regions (left end of each curve) inside the triangle (red), pentagon (green), and squares (blue) to the root of the tree (right end of each curve). Note that the length of these three branches is different (it explains that the root node appears at different abscissas.)

leaf-to-root order to construct the tree, followed by a second propagation in root-to-leaf order to canonize the tree. For further details, we refer the reader to [24, 5, 8]. During the tree construction, some useful information  $\mathcal{A}$  for computing the attribute function  $\mathcal{A}_{\lambda_s}$  can be updated incrementally: 1) Information accumulated on region, such as the area  $A$  and the sum of gray levels  $S_f$ ; 2) Information accumulated on contour, such as the contour length  $L$  and the sum of gradient's magnitude along the contour  $S_{\nabla}$ . Note that the computation of information on contour relies on the use of Khalimsky's grid proposed by [17]. A region boundary is composed of a set of elements materialized by 1-faces and 0-faces that lie between pixels (2-faces). An example is depicted in Fig. 6. We refer the reader to [37] for more details about these efficient information computations.

Once we have the tree structure  $\mathcal{T}$  represented by the parenthood image *parent*, and the corresponding information  $\mathcal{A}$  (e.g.,  $A$ ,  $S_f$ ,  $L$ , and  $S_{\nabla}$ ) for each node of the tree, we are able to compute the attribute function  $\mathcal{A}_{\lambda_s}$ . The computation of this attribute function  $\mathcal{A}_{\lambda_s}$  is detailed in Algorithm 1. We start by computing the initial values of attribute  $\mathcal{A}_{\lambda_s}$  according to Eq (6) by considering that only the underlying shape is removed (see line 11). Then we sort the shapes by increasing order of the average of gradient's magnitude along the shape contour  $\mathcal{A}_{\nabla}$ .

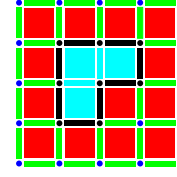


Figure 6: Materialization of pixels with 0-faces (blue disks), 1-faces (green strips), and 2-faces (red squares). The original pixels are the 2-faces, the boundaries are materialized with 0-faces and 1-faces. The contour of the cyan region is composed of black 1-faces and 0-faces.

We process the shapes in this order. For each underlying shape  $\tau$ , we compute a new value according to Eq (6) (see line 15), and update the value  $\mathcal{A}_{\lambda_s}(\tau)$  if the new value is greater. Then we remove the shape  $\tau$  from the tree and update the tree structure as well as the corresponding information.

The third step is to compute the extinction values  $\mathcal{E}$  for the local minima of the tree-based shape space weighted by the attribute  $\mathcal{A}_{\lambda_s}^{\downarrow}$ . This is achieved thanks to a Min-tree representation  $\mathcal{T}\mathcal{T}$  constructed on the tree-based shape space. The algorithm is described in Algorithm 2. The image *original\_min* tracks the smallest local minimum inside a lower level connected component  $CC$  of the tree-based shape space. For each local minimum shape  $\tau$ , the lowest  $CC$  that contains  $\tau$  and a smaller minimum is the lowest ancestor node whose smallest local minimum shape is different from  $\tau$ .

To compute the final saliency map  $\mathcal{M}_{\mathcal{E}}$  based on the extinction value, we weigh the boundaries (composed of 1-faces and 0-faces in the Khalimsky's grid) of each shape by the corresponding extinction value  $\mathcal{E}$ . More precisely, we weigh each 1-face  $e$  (resp. 0-face  $o$ ) by the maximal extinction value of the shapes whose boundaries contain  $e$  (resp.  $o$ ). The algorithm is given in Algorithm 3. It relies on two images *appear* and *vanish* defined on the 1-faces that are computed during the tree construction. The value *appear*( $e$ ) encodes the smallest region  $\mathcal{N}_a$  in the tree whose boundary contains the 1-face  $e$ , while *vanish*( $e$ ) denotes the smallest region  $\mathcal{N}_v$  that contains the 1-face  $e$  inside it.

## 4 Illustrations and experiments

In this section, we illustrate our proposed general framework with a simple segmentation model: piecewise-constant Mumford-Shah model. Using some more evolved energy functional will be one of our future work. For generic natural images, contours of significant objects usually coincide with segments of level lines. Our proposed method yields a hierarchical simplification rather than a hierarchical segmentation. So only qualitative illustrations are depicted in Section 4.1 for some images taken from the BSDS500 dataset introduced in [2]. For the Weizmann

```

1 COMPUTE_ATTRIBUTE(parent,  $\mathcal{T}$ ,  $\mathcal{A}$ )
2 for all  $\tau \in \mathcal{T}$  do
3    $A_R(\tau) \leftarrow A(\tau)$ ,  $S_{f,R}(\tau) \leftarrow S_f(\tau)$ ;
4   if  $\tau \neq \text{parent}(\tau)$  then  $Ch(\text{parent}(\tau)).\text{insert}(\tau)$ ;
5   for all  $\tau \in \mathcal{T}$  do
6     if  $\tau \neq \text{parent}(\tau)$  then
7        $A_R(\text{parent}(\tau)) \leftarrow A_R(\text{parent}(\tau)) - A(\tau)$ ;
8        $S_{f,R}(\text{parent}(\tau)) \leftarrow S_{f,R}(\text{parent}(\tau)) - S_f(\tau)$ ;
9   for all  $\tau \in \mathcal{T}$  do
10     $\mathcal{A}_\nabla(\tau) \leftarrow S_\nabla(\tau)/L(\tau)$ ,  $\tau_p \leftarrow \text{parent}(\tau)$ ;
11     $\mathcal{A}_{\lambda_s}(\tau) \leftarrow$ 
       $(\frac{S_{f,R}^2(\tau)}{A_R(\tau)} + \frac{S_{f,R}^2(\tau_p)}{A_R(\tau_p)} - \frac{(S_{f,R}(\tau_p) + S_{f,R}(\tau))^2}{A_R(\tau_p) + A_R(\tau)})/L(\tau)$ ;
12  $\mathcal{R}_\mathcal{T} \leftarrow \text{SORT\_NODES}(\mathcal{T}, \mathcal{A}_\nabla)$ 
13 for  $i \leftarrow 0$  to  $N_\mathcal{T}$  do
14    $\tau \leftarrow \mathcal{R}_\mathcal{T}(i)$ ,  $\tau_p \leftarrow \text{parent}(\tau)$ ;
15    $\lambda_t \leftarrow (\frac{S_{f,R}^2(\tau)}{A_R(\tau)} + \frac{S_{f,R}^2(\tau_p)}{A_R(\tau_p)} - \frac{(S_{f,R}(\tau_p) + S_{f,R}(\tau))^2}{A_R(\tau_p) + A_R(\tau)})/L(\tau)$ ;
16   if  $\lambda_t > \mathcal{A}_{\lambda_s}(\tau)$  then  $\mathcal{A}_{\lambda_s}(\tau) \leftarrow \lambda_t$ ;
17    $Ch(\tau_p).\text{remove}(\tau)$ ;
18   for all  $\tau_c \in Ch(\tau)$  do
19      $\text{parent}(\tau_c) \leftarrow \tau_p$ ;
20      $Ch(\tau_p).\text{insert}(\tau_c)$ ;
21    $A_R(\tau_p) \leftarrow A_R(\tau_p) + A(\tau)$ ;
22    $S_{f,R}(\tau_p) \leftarrow S_{f,R}(\tau_p) + S_f(\tau)$ ;
23 return  $\mathcal{A}_{\lambda_s}$ 

```

**Algorithm 1:** Computation of attribute function  $\mathcal{A}_{\lambda_s}$ . During the tree computation, we also compute four attribute information  $\mathcal{A}$ : region size  $A$ , region’s contour length  $L$ , sum of gray level  $S_f$  inside the region, and sum of gradient’s magnitude along the region contour  $S_\nabla$ .

segmentation database proposed by [1], the objects’ contours coincide with almost full level lines. Our method provides a hierarchical segmentation. Quantitative results using the associated evaluation framework are depicted in Section 4.2.

#### 4.1 Hierarchical color image pre-segmentation

In Fig. 7, we test our method on color images in the segmentation evaluation database proposed in [1]. Each image contains two objects to be segmented. We use the color tree of shapes proposed by [9], where the input image  $f$  in Eq (1) is a color image. A high parameter value  $\lambda_s = 8000$  is used, and the grain filter proposed in [21] is applied to get rid of too tiny shapes (*e.g.*, smaller than 10 pixels). Less than 100 level lines are selected, which results in a ratio of level line selection around 1157. These selected level lines form less than 200 regions in each image. The simplified images illustrated in Fig. 7 are obtained by taking the average color inside each region, where the boundaries between salient regions remain intact. Finding an actual segmentation becomes a lot easier with such a pre-segmentation. The extinction-based saliency maps  $\mathcal{M}_\mathcal{E}$  using the attribute  $\mathcal{A}_{\lambda_s}^\downarrow$  are depicted on the bottom of this figure. They represent hierarchical pre-segmentations.

Some illustrations of the extinction-based saliency map

```

1 COMPUTE_EXTINCTION_VALUE( $\mathcal{T}$ ,  $\mathcal{A}$ )
2 (parent $_\mathcal{T}$ ,  $\mathcal{R}_\mathcal{T}$ )  $\leftarrow$  COMPUTE_TREE( $\mathcal{A}$ );
3 for all  $\tau \in \mathcal{T}$  do  $\text{original\_min}(\tau) \leftarrow \text{undef}$ ;
4 for  $i \leftarrow 0$  to  $N_\mathcal{T}$  do
5    $\tau \leftarrow \mathcal{R}_\mathcal{T}(i)$ ,  $\tau_p \leftarrow \text{parent}_\mathcal{T}(\tau)$ ;
6   if  $\text{original\_min}(\tau) = \text{undef}$  then
7      $\text{original\_min}(\tau) = \tau$ ;
8   if  $\text{original\_min}(\tau_p) = \text{undef}$  then
9      $\text{original\_min}(\tau_p) \leftarrow \text{original\_min}(\tau)$ ;
10  else
11    if  $\mathcal{A}(\text{original\_min}(\tau_p)) > \mathcal{A}(\text{original\_min}(\tau))$ 
12      then
13         $\text{original\_min}(\tau_p) \leftarrow \text{original\_min}(\tau)$ ;
14  for all  $\tau \in \mathcal{T}$  do
15    if  $\tau$  is not a local minimum then
16       $\mathcal{E}(\tau) \leftarrow 0$ ;
17    else
18       $\tau_p \leftarrow \text{parent}_\mathcal{T}(\tau)$ ;
19      while
20         $\text{original\_min}(\tau_p) = \tau$  &  $\tau_p \neq \text{parent}_\mathcal{T}(\tau_p)$  do
21        |  $\tau_p \leftarrow (\text{parent}_\mathcal{T}(\tau_p))$ ;
22        |  $\mathcal{E}(\tau) \leftarrow \mathcal{A}(\tau_p) - \mathcal{A}(\tau)$ ;
23 return  $\mathcal{E}$ 

```

**Algorithm 2:** Computation of extinction values  $\mathcal{E}$  on a tree-based shape space weighted by an attribute  $\mathcal{A}$ . The image *original\_min* tracks the smallest local minimum shape inside a connected component of  $\mathcal{T}\mathcal{T}$ . Note that *parent<sub>T</sub>* encodes the Min-tree  $\mathcal{T}\mathcal{T}$  constructed on the shape space.

$\mathcal{M}_\mathcal{E}$  applied on images in the dataset of BSDS500 [2] are also shown in Fig. 8. Again, the input image  $f$  is a color image, and the color tree of shapes is used. As shown in Fig. 8, meaningful level lines are highlighted in  $\mathcal{M}_\mathcal{E}$  employing the attribute  $\mathcal{A}_{\lambda_s}^\downarrow$ . Hierarchical image simplification results can then be obtained by thresholding  $\mathcal{M}_\mathcal{E}$ . The saliency maps for all the images in the dataset of BSDS500 is available on <http://publications.lrde.epita.fr/xu.hierarchyms11>.

We have also tested our method on some cellular images, where the method is applied on the color input image  $f$ . As illustrated in Fig. 9, the cellular image is strongly simplified, which almost leads to a uniform background. Finding an actual cellular segmentation result would become much easier.

#### 4.2 Evaluation in context of segmentation

We have also evaluated our hierarchical image simplifications in context of segmentation on Weizmann segmentation evaluation database in [1]. For the 100 images containing 2 objects in this database (See Fig. 7 for several examples), the saliency maps are thresholded with a fixed thresholding value to yield a partition result. And we filter out the regions whose area is less than 100 pixels. Note that, in order to perform a fair comparison with the state-of-the art, the saliency maps are constructed using grayscale tree of shapes computed by [15] on grayscale versions of the input images  $f$ . We performed two tests as



```

1 COMPUTE_SALIENCY_MAP( $f$ )
2 ( $parent, \mathcal{T}, \mathcal{A}$ )  $\leftarrow$  COMPUTE_TREE( $f$ );
3  $\mathcal{A}_{\lambda_s} \leftarrow$  COMPUTE_ATTRIBUTE( $parent, \mathcal{T}, \mathcal{A}$ );
4  $\lambda_s^M \leftarrow 0$ ;
5 for all  $\tau \in \mathcal{T}$  do  $\lambda_s^M \leftarrow \max(\lambda_s^M, \mathcal{A}_{\lambda_s}(\tau))$ ;
6 for all  $\tau \in \mathcal{T}$  do  $\mathcal{A}_{\lambda_s}^\downarrow \leftarrow \lambda_s^M - \mathcal{A}_{\lambda_s}(\tau)$ ;
7  $\mathcal{E} \leftarrow$  COMPUTE_EXTINCTION_VALUE( $\mathcal{T}, \mathcal{A}_{\lambda_s}^\downarrow$ );
8 for all 1-face  $e$  do  $\mathcal{M}_\mathcal{E}(e) \leftarrow 0$ ;
9 for all  $e$  do
10    $N_a \leftarrow appear(e), N_v \leftarrow vanish(e)$ ;
11   while  $N_a \neq N_v$  do
12      $\mathcal{M}_\mathcal{E}(e) \leftarrow \max(\mathcal{E}(N_a), \mathcal{M}_\mathcal{E}(e))$ ,
13      $N_a \leftarrow parent(N_a)$ ;
14 for all 0-face  $o$  do
15    $\mathcal{M}_\mathcal{E}(o) \leftarrow \max(\mathcal{M}_\mathcal{E}(e_1), \mathcal{M}_\mathcal{E}(e_2), \mathcal{M}_\mathcal{E}(e_3), \mathcal{M}_\mathcal{E}(e_4))$ ;
16 return  $\mathcal{M}_\mathcal{E}$ 

```

**Algorithm 3:** Computation of saliency map  $\mathcal{M}_\mathcal{E}$  representing a hierarchical result of level line selections. The 1-faces  $e_1, e_2, e_3, e_4$  are the 1-faces adjacent to  $o$ .

presented in [1] based on F-measure and number of fragments. For a segmentation  $Seg$  and a ground truth segmentation of the object  $GT$ , the F-measure is defined by  $F = 2 \times precision \times recall / (precision + recall)$ , where  $precision = |Seg \cap GT| / |Seg|$ ,  $recall = |Seg \cap GT| / |GT|$ . The number of fragments is the number of regions selected from a partition to form the object segmentation result  $Seg$ .

In the first test, for each foreground object, we select the segment that fits it the best based on F-measure score. The results of this single segment coverage test is depicted in Table 1 (See [1] for implementation details on the settings for the other methods). In this test, our method achieves F-measure score on par with the state-of-the-art methods, especially when replacing, in the attribute  $\mathcal{A}_\nabla$ , the classical grayscale gradient with the (learned) gradient computed (on the grayscale input image  $f$ ) by [12] (see “Our2”). By using another gradient, we change the order in which the nodes of the tree are processed; thus this result highlights the importance of the sorting step in our algorithm. In Table 1, note that *Gpb-owt-ucm without texture* denotes the method of Gpb-owt-ucm computed without taking into account the texture information in the Gpb part. More precisely, in this case, the Gpb is computed using only brightness and color gradients. Note also that our method does not explicitly use any texture information either.

In the second test, a combination of segments whose area overlaps considerably the foreground objects is utilized to assess the performance. For each union of segments, we measure the F-measure score and the number of segments composing it. This test is a compromise between good F-measure and low number of fragments. The results of this fragmentation test is given in Table 2. In this test the averaged F-measure of different methods is fairly sim-



Figure 7: Some pre-segmentation results obtained with our proposed method on the segmentation evaluation database in [1]. Top: input images; Middle: pre-segmentations obtained with the simplification method; Bottom: inverted saliency maps for hierarchical simplifications.

ilar. However, our method, as a pre-segmentation method without using any texture information has a relatively high number of fragments.

The saliency maps for these images containing two objects in the Weizmann dataset is available on <http://publications.lrde.epita.fr/xu.hierarchysll>.

Table 1: Results of single segment coverage test using F-measure on the two objects dataset in [1]. “Our2” stands for our method with the attribute  $\mathcal{A}_\nabla$  using the gradient computed by [12].

Method	Average	Larger	Smaller
<i>Our2</i>	<b>0.80</b>	<b>0.79</b>	<b>0.81</b>
Gpb-owt-ucm	<b>0.79</b>	<b>0.80</b>	<b>0.78</b>
<i>Our</i>	0.77	0.77	0.76
Gpb-owt-ucm without texture	0.76	0.76	0.75
Gpb in [1]	0.72	0.70	0.75
Method in [1]	0.68	0.70	0.65
SWA in [31]	0.66	0.74	0.57
MeanShift	0.61	0.65	0.58
N-Cuts	0.58	0.66	0.49

## 5 Comparison with similar works

The tree of shapes has been widely used in connected operators, filtering tools that act by merging flat zones for image simplification and segmentation. The simplification and segmentation relies on relevant shapes extraction (*i.e.*, salient level lines), usually achieved by tree filtering based on some attribute function. A detailed review of tree filtering strategies can be found in [30]. In all these strategies,

Table 2: Fragmented coverage test results for the two objects dataset proposed by [1]: compromise between good F-measure and low number of fragments. Our results are comparable to the state-of-the-art. “Our2” stands for our method with the attribute  $\mathcal{A}_\nabla$  using the gradient computed by [12].

Method	Averaged		Larger object		Smaller object	
	F-measure	#fragments	F-measure	#fragments	F-measure	#fragments
SWA in [31]	<b>0.88</b>	3.13	<b>0.91</b>	3.88	<b>0.84</b>	2.37
<i>Our2</i>	<b>0.86</b>	2.40	0.85	3.00	<b>0.86</b>	1.81
Method in [1]	0.85	<b>1.67</b>	0.87	<b>2.00</b>	<b>0.84</b>	<b>1.33</b>
N-Cuts in [32]	0.84	2.64	<b>0.88</b>	3.34	0.80	1.93
Gpb reported in [1]	0.84	2.95	0.87	3.60	0.81	2.30
<i>Our</i>	0.83	3.16	0.85	4.10	0.81	2.23
Gpb-owt-ucm in [2]	0.82	<b>1.57</b>	0.84	<b>1.79</b>	0.81	<b>1.35</b>
Gpb-owt-ucm without texture	0.81	2.72	0.82	3.32	0.80	2.12
MeanShift in [11]	0.78	3.65	0.85	4.49	0.71	2.81

the attribute function  $\mathcal{A}$  characterizing each node plays a very important role in connected filtering. The classical connected operators make filtering decisions based only on attribute function itself or the inclusion relationship of the tree (*e.g.*, [40]). They are usually performed by removing the nodes whose attributes are lower than a given threshold. The method we propose in this paper combines this idea of classical connected operators with the energy minimization problem of Eq (2). It also makes use of the spatial information of the original image from which the tree is constructed. This might give more robust filtering decision.

In this paper, we focus particularly on hierarchical relevant shapes selection by minimizing some multiscale affine separable energy functional (*e.g.*, piecewise-constant Mumford-Shah functional). The closely related work is the one in [16], where the authors proposed the scale-set theory, including an efficient greedy algorithm to minimize this kind of energy on a hierarchy of segmentations. More precisely, the authors use dynamic programming to efficiently compute two scale parameters  $\lambda_s^+$  and  $\lambda_s^-$  for each region  $R$  of the input hierarchy  $H$ , where  $\lambda_s^+$  (*resp.*,  $\lambda_s^-$ ) corresponds to the smallest parameter  $\lambda_s$  such that the region  $R \in H$  belongs (*resp.*, does not belong) to the optimal solution of segmentation by minimizing  $E_{\lambda_s}$ , we have  $\lambda_s^-(R) = \min_{R' \in H, R \subset R'} \lambda_s^+(R')$ . There may exist some regions  $R$  such that  $\lambda_s^-(R) \leq \lambda_s^+(R)$ , which implies that the region  $R \in H$  does not belong to any optimal cut of  $H$  by minimizing the energy  $E_{\lambda_s}$ . One removes these regions from the hierarchy  $H$  and updates the parenthood relationship which yields a hierarchy  $H'$ , a hierarchy of global optimal segmentations on the input hierarchy. This work has been continued and extended by [18]. These methods work on an input hierarchy of segmentations, which is very different from the tree of shapes (a natural and equivalent image representation). In this sense, our method can be seen as the extension of scale-set theory to the tree of shapes.

Another related work is the one in [4]. It also selects meaningful level lines for image simplification and segmen-

tation by minimizing the piecewise-constant Mumford-Shah functional. For this method, at each step the level line is selected which inflicts the largest decrease of functional. As a consequence, the iterative process of [4] requires not only computing a lot of information to be able to update the functional after each level line suppression, but also to find at each step, among all remaining level lines, the one candidate to the next removal. Consequently, the optimization process has a  $O(N_\mathcal{T}^2)$  time complexity w.r.t. the number of nodes  $N_\mathcal{T}$  of the tree. A heap-based implementation may improve the time complexity, but since at each removal, one has to update the corresponding energy variation for its children, parent, siblings, maintaining the heap structure is a costly process. In practice, the gain using heap-based implementation is relatively insignificant. Hence [4] is computationally expensive. We propose to fix that issue thanks to a reasonable ordering of level lines based on their quantitative meaningfulness measurement  $\mathcal{A}$ . The time complexity of our optimization process is linear w.r.t. the number of nodes  $N_\mathcal{T}$ . We have implemented the method of [4] using the same tree construction algorithm and the same data structure based on heap. We have compared the running time on 7 classic images on a regular PC station. The comparison is detailed in Table. 3. Our proposal is significantly faster than that of [4]. Our approach is almost linear w.r.t. the number of nodes in the tree. Yet, the method of [4] seems to depend also on the depth of the tree. In [4], the authors proposed applying the simplification scheme successively with a set of augmenting parameters  $\lambda_s$  so that to construct the input hierarchy. Then they employed the scheme of scale-set theory proposed by [16] on the obtained hierarchy to achieve a final hierarchy of optimal segmentations. In our case, rather than using a fixed parameter  $\lambda_s$  or a set of fixed parameters, we propose to assign a measure related to  $\lambda_s$  to each shape as an attribute function. Then we use the hierarchy transformation (reviewed in Section 2.3) based on extinction values and on a tree-based shape space to compute a hierarchical salient level lines selection.

It is worth noticing that the minimization of Mumford-

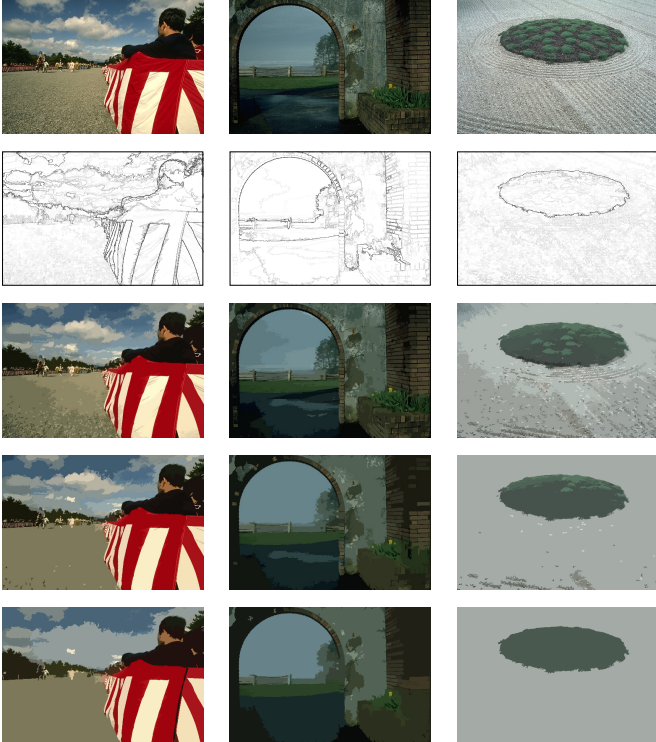


Figure 8: Illustration of the hierarchical image simplification using the attribute function  $\mathcal{A}_{\lambda_s}^\downarrow$ , applied on some images from the dataset of BSDS500 proposed by [2]. From top to bottom: input images; inverted saliency maps; slight simplification; moderate simplification; strong simplification.

Shah-like functional has also been applied to shape analysis in [33]. It consists of adding a non-local term, which is the squared average of the field in the energy functional. Its minimization tends to form negative field values on narrow or small parts as well as on protrusions, and positive field values on central part(s) of the input shape. The negative and positive regions inside the input shape yield some saddle points at which a crossing of a level curve occurs. This leads to a binary partition hierarchy  $H_b$  of the shape. Then a probability measure based on the obtained field values inside the shape is assigned to each node of the partition hierarchy  $H_b$ . A set of hierarchical representations of the shape is obtained by removing some nodes from  $H_b$  and update the parenthood relationship. Each candidate hierarchical representation is assigned with a saliency measure given by the products of the probability measure of the removed nodes. These hierarchical representations of the shape associated with the global saliency are used to analyze the shape. Our proposal is different from this framework in terms of the use of energy minimization. In [33], the energy minimization is used to create an image with negative and positive regions for a given shape. Then one constructs a binary hierarchy of partitions of the created image via its saddle points, and weighs each node based on the obtained image values. In our case,

Table 3: Comparison of computation times on seven classical images. The size for image “House” and “Camera” is  $256 \times 256$ , and  $512 \times 512$  for the other images.

Image	Depth	#Nodes	Time (s)	
			[4]	Our
<i>House</i>	126	23588	4.11	0.22
<i>Camera</i>	126	24150	4.19	0.23
<i>Lena</i>	141	84699	27.77	0.92
<i>Peppers</i>	176	97934	48.18	0.93
<i>Boat</i>	255	100518	87.24	0.94
<i>Barbara</i>	131	106285	51.87	0.99
<i>Mandrill</i>	185	153029	200.22	1.34

the energy minimization is performed on an input image subordinated to its hierarchical representation by the tree of shapes. This yields a quantitative meaningfulness measure for each node of the tree of shapes.

## 6 Conclusion

In this paper, we have presented an efficient approach of hierarchical image simplification and segmentation based on minimizing some multiscale separable energy functional on the tree of shapes, a unique and equivalent image representation. It relies on the idea of hierarchy transformation based on extinction values and on a tree-based shape space to compute a saliency map representing the final hierarchical image simplification and segmentation. The salient structures in images are highlighted in this saliency map. A simplified image with preservation of salient structures can be obtained by thresholding the saliency map. Some qualitative illustrations and quantitative evaluation in context of image segmentation on a public segmentation dataset demonstrate the efficiency of the proposed method. A binary executable of the proposed approach is available on <http://publications.lrde.epita.fr/xu.hierarchyms11>.

In the future, we would like to explore some applications employing a strongly simplified image as pre-processing step. We believe that this could be useful for analyzing high-resolution satellite images and images with texts, where the contours of meaningful objects in images usually coincide with full level lines. Besides, as advocated in Table 1 for Gpb-owt-ucm, the texture provides important information for image segmentation. An interesting perspective is to incorporate texture information in our proposed framework. Since the tree of shapes is a natural representation of the input image, a possible way to integrate texture information might consist in replacing the original image with a new grayscale image incorporating texture features. Although this is not directly applicable to our case, probability map incorporating region features, such as proposed in [3], are worth exploring. Computing the tree of shapes of such probability maps has already

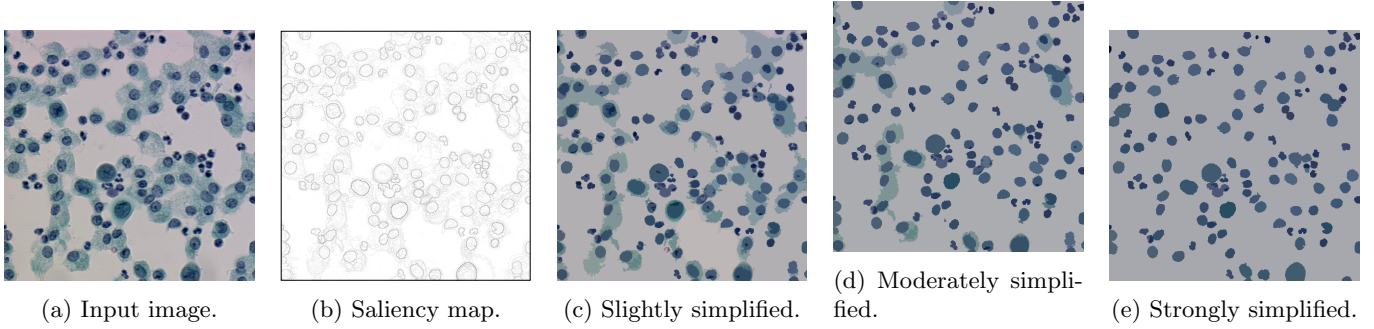


Figure 9: Illustration of the proposed hierarchical pre-segmentations on a cellular image.

been proved valuable (see [13]). In another direction, it would be interesting to investigate some other energy functionals for some specific tasks. Examples are the rate distortion optimization used in image or video compression coding system (see [29, 4]) and the energy based on spectral unmixing used for hyperspectral image segmentation in [36]. The energy functionals in these works are affine separable, which straightforwardly allows to use them in our proposed framework. Last, but not the least, given that using a learned gradient improves the results, a major research avenue is to combine our approach with learning techniques.

## References

- [1] S. Alpert, M. Galun, A. Brandt, and R. Basri. Image segmentation by probabilistic bottom-up aggregation and cue integration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(2):315–327, 2012.
- [2] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011.
- [3] X. Bai and G. Sapiro. Geodesic matting: A framework for fast interactive image and video segmentation and matting. *International Journal of Computer Vision*, 82(2):113–132, 2009.
- [4] C. Ballester, V. Caselles, L. Igual, and L. Garrido. Level lines selection with variational models for segmentation and encoding. *Journal of Mathematical Imaging and Vision*, 27:5–27, 2007.
- [5] C. Berger, T. Géraud, R. Levillain, N. Widynski, A. Baillard, and E. Bertin. Effective component tree computation with application to pattern recognition in astronomical imaging. In *IEEE International Conference on Image Processing*, volume 4, pages 41–44, 2007.
- [6] F. Cao, P. Musé, and F. Sur. Extracting meaningful curves from images. *Journal of Mathematical Imaging and Vision*, 22(2-3):159–181, 2005.
- [7] J. Cardelino, G. Randall, M. Bertalmio, and V. Caselles. Region based segmentation using the tree of shapes. In *IEEE International Conference on Image Processing*, pages 2421–2424, 2006.
- [8] E. Carlinet and T. Géraud. A comparative review of component tree computation algorithms. *IEEE Transactions on Image Processing*, 23(9):3885–3895, Sept. 2014.
- [9] E. Carlinet and T. Géraud. MToS: A tree of shapes for multivariate images. *IEEE Transactions on Image Processing*, 24(12):5330–5342, 2015.
- [10] V. Caselles, B. Coll, and J.-M. Morel. Topographic maps and local contrast changes in natural images. *International Journal of Computer Vision*, 33(1):5–27, 1999.
- [11] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.
- [12] P. Dollár and C. L. Zitnick. Fast edge detection using structured forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8):1558–1570, 2015.
- [13] A. Dubrovina, R. Hershkovitz, and R. Kimmel. Image editing using level set trees. In *IEEE International Conference on Image Processing*, pages 4442–4446, 2014.
- [14] P. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, sep 2004.
- [15] T. Géraud, E. Carlinet, S. Crozet, and L. Najman. A quasi-linear algorithm to compute the tree of shapes of  $nD$  images. In *International Symposium on Mathematical Morphology*, pages 98–110, 2013.
- [16] L. Guigues, J. P. Cocquerez, and H. Le Men. Scale-sets image analysis. *International Journal of Computer Vision*, 68(3):289–317, 2006.
- [17] E. Khalimsky, R. Kopperman, and P. R Meyer. Computer graphics and connected topologies on finite or-

- dered sets. *Topology and its Applications*, 36(1):1–17, 1990.
- [18] B. R. Kiran and J. Serra. Global-local optimizations by hierarchical cuts and climbing energies. *Pattern Recognition*, 47(1):12–24, 2014.
- [19] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50(5):363–370, 1984.
- [20] H. Lu, J. C. Woods, and M. Ghanbari. Binary partition tree analysis based on region evolution and its application to tree simplification. *IEEE Transactions on Image Processing*, 16(4):1131–1138, 2007.
- [21] P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Transactions on Image Processing*, 9(5):860–872, 2000.
- [22] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [23] L. Najman. On the equivalence between hierarchical segmentations and ultrametric watersheds. *Journal of Mathematical Imaging and Vision*, 40:231–247, 2011.
- [24] L. Najman and M. Couprie. Building the component tree in quasi-linear time. *IEEE Transactions on Image Processing*, 15(11):3531–3539, 2006.
- [25] L. Najman and M. Schmitt. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12):1163–1173, 1996.
- [26] A. Pardo. Semantic image segmentation using morphological tools. In *IEEE International Conference on Image Processing*, pages 745–748, 2002.
- [27] B. Perret, J. Cousty, J. C. R. Ura, and S. J. F. Guimarães. Evaluation of morphological hierarchies for supervised segmentation. In *International Symposium on Mathematical Morphology*, pages 39–50. Springer, 2015.
- [28] C. Ronse. Ordering partial partitions for image segmentation and filtering: Merging, creating and inflating blocks. *Journal of Mathematical Imaging and Vision*, 49(1):202–233, 2014.
- [29] P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation and information retrieval. *IEEE Transactions on Image Processing*, 9(4):561–576, 2000.
- [30] P. Salembier and M. H. Wilkinson. Connected operators. *Signal Processing Magazine*, 26(6):136–157, 2009.
- [31] E. Sharon, M. Galun, D. Sharon, R. Basri, and A. Brandt. Hierarchy and adaptivity in segmenting visual scenes. *Nature*, 442(7104):810–813, 2006.
- [32] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.
- [33] S. Tari and M. Genctav. From a non-local ambrosio-tortorelli phase field to a randomized part hierarchy tree. *Journal of mathematical imaging and vision*, 49(1):69–86, 2014.
- [34] R. Tarjan. Efficiency of a good but not linear set union algorithm. *Journal of the ACM*, 22(2):215–225, 1975.
- [35] C. Vachier and F. Meyer. Extinction value: a new measurement of persistence. In *IEEE Workshop on Nonlinear Signal and Image Processing*, volume 1, pages 254–257, 1995.
- [36] M. A. Veganzones, G. Tochon, M. Dalla-Mura, A. J. Plaza, and J. Chanussot. Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation. *IEEE Transactions on Image Processing*, 23(8):3574–3589, 2014.
- [37] Y. Xu, E. Carlinet, T. Géraud, and L. Najman. Efficient computation of attributes and saliency maps on tree-based image representations. In *International Symposium on Mathematical Morphology*, volume 9082 of *LNCS*, pages 693–704. Springer, 2015.
- [38] Y. Xu, T. Géraud, and L. Najman. Salient level lines selection using the mumford-shah functional. In *IEEE International Conference on Image Processing*, pages 1227–1231, 2013.
- [39] Y. Xu, T. Géraud, and L. Najman. Two applications of shape-based morphology: Blood vessels segmentation and a generalization of constrained connectivity. In *International Symposium on Mathematical Morphology*, pages 390–401, 2013.
- [40] Y. Xu, T. Géraud, and L. Najman. Connected filtering on tree-based shape-spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015. to appear.