Pattern Recognition Letters
journal homepage: www.elsevier.com

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

Yongchao Xu[a,b,c,**], Thierry Géraud[a], Laurent Najman[c]

[a] *EPITA Research and Development Laboratory (LRDE), 14-16, rue Voltaire, FR-94276 Le Kremlin-Bicêtre, France*
[b] *Department of Signal and Image Processing, Télécom ParisTech, 46 rue Barrault, 75013 Paris, France*
[c] *Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge (LIGM), A3SI, ESIEE Paris, Cité Descartes, BP 99 FR-93162 Noisy-le-Grand, France*

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 the piecewise-constant Mumford-Shah 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.*, Caselles et al. (1999); Cao et al. (2005)) 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 Monasse and Guichard (2000), and also known as *topographic map* in Caselles et al. (1999). 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 Pardo (2002); Cao et al. (2005); Cardelino et al. (2006). In Lu et al. (2007), 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 λ_s is a parameter. A popular example is the seminal work of Mumford and Shah (1989). Curve evolution methods are usually used to solve this minimization problem. They have solid theoretical foundations, yet they are often computational 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 Najman and Schmitt (1996) and gpb-owt-ucm proposed by Arbelaez et al. (2011) 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 segmen-

**Corresponding author: Tel.: +33 1 53 14 59 43; fax: +33 1 53 14 59 13;
e-mail: yongchao.xu@lrde.epita.fr (Yongchao Xu)

tations in the sense of energy minimization. Examples are the works of Guigues et al. (2006); Kiran and Serra (2014). Yet, the choice or the construction of the input hierarchy of segmentations for these methods is an interesting problem in itself. Perret et al. (2015) 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 Xu et al. (2013a), in which we proposed an efficient greedy algorithm computing a locally optimal solution of the energy minimization problem. The basic idea is to take into account the meaningfulness of each level line which measures its "importance". An example of meaningfulness function that we will use through the paper is the average of gradient's magnitude along level lines. The order based on these meaningfulness values allows to get very quickly a locally optimal solution, which yields a well-simplified image while preserving the salient structures. The current paper extends this idea to hierarchical simplification and segmentation. More precisely, following the same principle but without fixing the parameter λ_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 representing the complete hierarchical simplifications or segmentations. To do so, we rely on the idea of hierarchy transformation via extinction value proposed by Vachier and Meyer (1995) and on the framework of tree-based shape space introduced in Xu et al. (2015b). This scheme of hierarchy transformation has been first used in Xu et al. (2013b) for a different input hierarchy and attribute function. Related algorithms were presented in Xu et al. (2015a). The present paper extends on these ideas, focusing on the computation of an attribute function related to energy minimization.

The main contribution of this current paper is the proposition of a general framework of hierarchical image simplification and segmentation method based on energy minimization subordinated to the tree of shapes, contrary to the classical approaches that are subordinated to an initial hierarchy of segmentations. It is based on the introduction of a novel attribute function $\mathcal{A}_{\lambda_s}$ related to energy minimization. We have tested the proposed framework 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 Alpert et al. (2012). In particular, they are on par with Gpb-owt-ucm proposed in Arbelaez et al. (2011) 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.

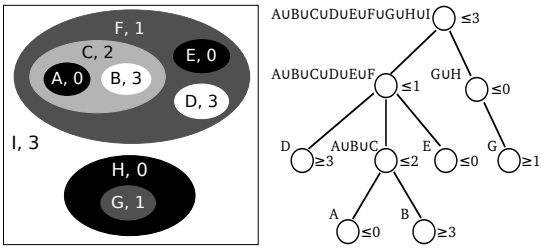


Fig. 1: An image (left) and its tree of shapes (right).

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 $f : \Omega \to \mathbb{R}$ 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 Monasse and Guichard (2000) 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 Carlinet and Géraud (2015) 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: $H = \{\mathcal{P}_i \mid 0 \leq i \leq n, \forall j,k,\ 0 \leq j \leq k \leq n \Rightarrow \mathcal{P}_j \sqsubseteq \mathcal{P}_k\}$, where $\mathcal{P}_n$ is the partition $\{\Omega\}$ of Ω into a single region, and $\mathcal{P}_0$ represents the finest partition of the image f. $\mathcal{P}_j \sqsubseteq \mathcal{P}_k$ 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 Koenderink (1984). From this principle, for any couple of scales $\lambda_{s_2} > \lambda_{s_1}$, the "structures" found at scale λ_{s_2} should find a "cause" at scale λ_{s_1}. In the case of a hierarchy of segmentation, following the work of Guigues et al. (2006), 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, λ_s) is called an indexed hierarchy.

A useful representation of hierarchical image segmentations was originally introduced in Najman and Schmitt (1996) 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 Guigues et al. (2006) through the notion of scale-set theory for visualization purposes, and it is then popularized by Arbelaez et al. (2011) under the name of *ultrametric contour map* for boundary extraction and comparing hierarchies. It has been proved theoretically in Najman (2011) that a hierarchy of segmentations is equivalent to a saliency map. Roughly speaking, for a given indexed hierarchy (H, λ), the corresponding saliency map can be obtained by weighing each contour of the image domain with the highest value λ_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 Xu et al. (2015b): 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 Xu et al. (2015b), 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 attribute $\mathcal{A}$. 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 Vachier and Meyer (1995). Let $\prec$ be a strict total order on the set of minima $m_1 \prec m_2 \prec \ldots$, 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 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 Xu et al. (2013b),

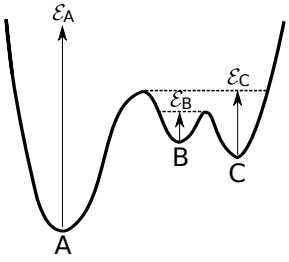


Fig. 2: Illustration of the extinction values $\mathcal{E}$ of three minima. The order is $A \prec C \prec B$. B merges with C, C merges with A.

where the input hierarchy is a minimum spanning tree and the attribute is computed locally inspired from the work of Felzenszwalb and Huttenlocher (2004). For the sake of completeness, the algorithm Xu et al. (2015a) 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 Guigues et al. (2006) and Kiran and Serra (2014)). 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 λ_s is a parameter. Let $\{R\} = R_1 \sqcup \cdots \sqcup R_n$ 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_{λ_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_{λ_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 Mumford and Shah (1989). For an image f, it is given by

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

where $\tilde{f}_i = \frac{1}{|R_i|} \iint_{R_i} f \, dxdy$ 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 is a general framework of hierarchical image simplification and segmentation based on energy minimization subordinated to the tree of shapes. It extends a preliminary version of this study in Xu et al. (2013a) that selects salient level lines based on Mumford-Shah energy functional minimization. We review this non-hierarchical version in Section 3.2, using a more general multiscale affine separable energy. The hierarchical version proposed in the current paper is achieved thanks to:

- the introduction of a novel attribute function $\mathcal{A}_{\lambda_s}$ for each level line related to the energy regularization parameter λ_s,

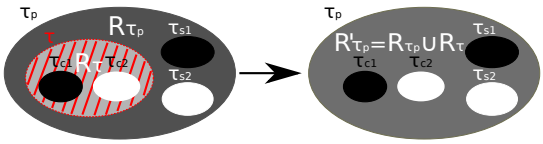


Fig. 3: Suppressing the node τ makes the "region" R_τ (covered with red oblique lines) merge with R_{τ_p}; the result (depicted in the right image) is a simplified image.

- and the idea of hierarchy transformation based on extinction values and on a tree-based shape space.

This is detailed in Section 3.3. Section 3.4 provides 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 $\{R_\mathcal{T}\} = R_{\tau_1} \sqcup \cdots \sqcup R_{\tau_n}$, where $R_\tau = \{p \,|\, p \in \tau, p \notin Ch(\tau)\}$ with $Ch(\tau)$ representing all the children of the shape τ. 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:

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

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 τ. The parent of its children $\tau_{c1}, \ldots, \tau_{ck}$ should also be updated to the τ_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 Ballester et al. (2007)). It iteratively removes the shapes to decrease the energy functional. The greedy algorithm stops when no other shape can be removed that favors a decrease of the energy. The removability of a shape τ is decided by the sign of the energy variation $\Delta E^\tau_{\lambda_s}$ while τ is suppressed. For the multiscale affine separable energy described in Section 2.4, $\Delta E^\tau_{\lambda_s}$ is given by:

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

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 $\Delta E^\tau_{\lambda_s}$ is given by:

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

Fig. 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.

If $\Delta E^\tau_{\lambda_s}$ is negative, which means the suppression of τ decreases the functional, then we remove τ. According to Eq. (3), the removability of a shape τ depends only on R_τ and R_{τ_p}. As the suppression of the shape τ triggers the update of R_{τ_p}, the removal of τ impacts also the removability of its parent, its children and siblings. So the order of level line removal is critical. In Xu et al. (2013a), we proposed to fix the order by sorting the level lines in increasing order of a quantitative meaningfulness attribute $\mathcal{A}$ (*e.g.*, the 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) attribute $\mathcal{A}_\nabla$ are easier (*resp.* more difficult) to filter out under the energy minimization process of Eq. (2). Consequently, the level line sorting based on attribute $\mathcal{A}_\nabla$ provides a reasonable order to perform the level lines suppression that makes the energy functional decrease. Indeed, initially, each region R_τ 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 λ_s, in the multiscale affine separable energy, controls the simplification/segmentation degree for the method described in Section 3.2, which is however not hierarchical. Because some level line τ may be removed with a parameter λ_{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 Ballester et al. (2007) is not hierarchical either.

Instead of fixing the parameter λ_s in the energy functional (*e.g.*, λ_s in Eq (1)), we propose to compute an individual λ_s for each shape of the tree following the same principle of the energy minimization. For a given λ_s, the removability of a shape τ is based on the sign of energy variation $\Delta E^\tau_{\lambda_s}$ in Eq. (3), which is a linear decreasing function w.r.t. λ_s (*e.g.*, the Eq (4) for the piecewise-constant Mumford-Shah functional). When λ_s is bigger than some value $\lambda_{s_{min}}$, $\Delta E^\tau_{\lambda_s}$ 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 τ, the attribute function $\mathcal{A}_{\lambda_s}(\tau)$ defined in Eq (5) depends on R_τ and R_{τ_p}, which means $\mathcal{A}_{\lambda_s}(\tau)$ is decided by the shape τ itself, its parent, its siblings, and its children. Because the attribute function $\mathcal{A}_{\lambda_s}$ is computed under the hypothesis that the shape τ under scrutiny is suppressed, we also need to update R_{τ_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 τ. 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:

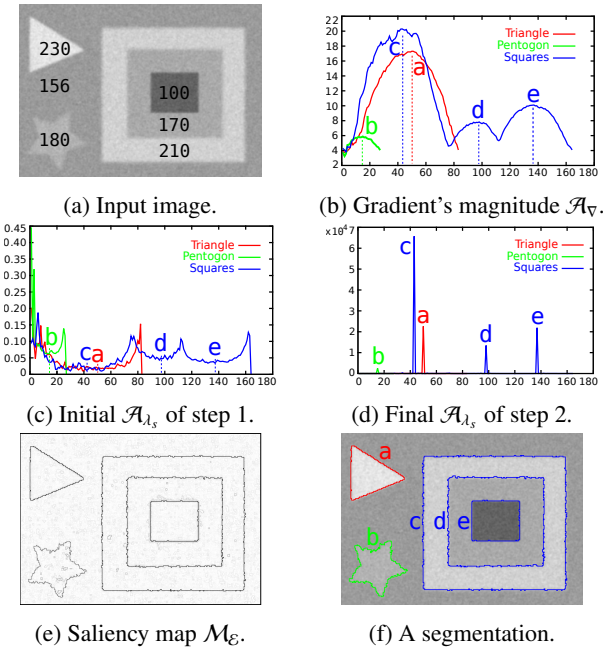
step 1: Compute $\mathcal{A}_{\lambda_s}$ for each shape $\tau \in \mathcal{T}$ supposing that only the shape under scrutiny is removed, and sort the set of shapes $\{\tau \,|\, \tau \in \mathcal{T}\}$ in increasing order of shape meaningfulness indicated by an attribute $\mathcal{A}$ (*e.g.*, $\mathcal{A}_\nabla$).

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 τ, and update it if the value is greater than the older one. Update also the parenthood relationship and the corresponding information for R_{τ_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 τ. 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: $\mathcal{A}^{\downarrow}_{\lambda_s}(\tau) = \max_{\tau' \in \mathcal{T}}(\mathcal{A}_{\lambda_s}(\tau')) - \mathcal{A}_{\lambda_s}(\tau)$. 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



(a) Input image. (b) Gradient's magnitude $\mathcal{A}_\nabla$.

(c) Initial $\mathcal{A}_{\lambda_s}$ of step 1. (d) Final $\mathcal{A}_{\lambda_s}$ of step 2.

(e) Saliency map $\mathcal{M}_\mathcal{E}$. (f) A segmentation.

Fig. 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.)

not effective to measure the importance of the shapes: indeed, this is due to the very small size of each region in $\{R_\mathcal{T}\}$. 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 obtaiend by thresholding $\mathcal{M}_\mathcal{E}$ are depicted in Fig. 5. (e) and (f).

3.4. Implementation

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

Once we have the tree structure $\mathcal{T}$ represented by the parenthood image *parent*, and the corresponding information $\mathcal{A}$ for each node of the tree, we are able to compute the attribute function $\mathcal{A}_{\lambda_s}$. Note that the information $\mathcal{A}$ can be the area A, the sum of gray levels S_f, the contour length L, or the sum of gradient's magnitude along the contour S_∇. The computation of this attribute function $\mathcal{A}_{\lambda_s}$ is performed while computing the tree, and 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 $\mathcal{A}_\nabla$. We process the shapes in this order. For each underlying shape τ, 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

```
COMPUTE_ATTRIBUTE(parent, 𝒯, 𝒜)
for all τ ∈ 𝒯 do
    A_R(τ) ← A(τ), S_{f,R}(τ) ← S_f(τ);
    if τ ≠ parent(τ) then Ch(parent(τ)).insert(τ);
for all τ ∈ 𝒯 do
    if τ ≠ parent(τ) then
        A_R(parent(τ)) ← A_R(parent(τ)) − A(τ);
        S_{f,R}(parent(τ)) ← S_{f,R}(parent(τ)) − S_f(τ);
for all τ ∈ 𝒯 do
    𝒜_∇(τ) ← S_∇(τ)/L(τ), τ_p ← parent(τ);
    𝒜_{λ_s}(τ) ← (S²_{f,R}(τ)/A_R(τ) + S²_{f,R}(τ_p)/A_R(τ_p) − (S_{f,R}(τ_p)+S_{f,R}(τ))²/(A_R(τ_p)+A_R(τ)))/L(τ);
ℛ_𝒯 ← SORT_NODES(𝒯, 𝒜_∇)
for i ← 0 to N_𝒯 do
    τ ← R_𝒯(i), τ_p ← parent(τ);
    λ_t ← (S²_{f,R}(τ)/A_R(τ) + S²_{f,R}(τ_p)/A_R(τ_p) − (S_{f,R}(τ_p)+S_{f,R}(τ))²/(A_R(τ_p)+A_R(τ)))/L(τ);
    if λ_t > 𝒜_{λ_s}(τ) then 𝒜_{λ_s}(τ) ← λ_t;
    Ch(τ_p).remove(τ);
    for all τ_c ∈ Ch(τ) do
        parent(τ_c) ← τ_p;
        Ch(τ_p).insert(τ_c);
    A_R(τ_p) ← A_R(τ_p) + A_R(τ);
    S_{f,R}(τ_p) ← S_{f,R}(τ_p) + S_{f,R}(τ);
return 𝒜_{λ_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_∇.

the shape τ from the tree and update the tree structure as well as the corresponding information.

The second step is to compute the extinction values $\mathcal{E}$ for the local minima of the tree-based shape space weighted by the attribute $\mathcal{A}^{\downarrow}_{\lambda_s}$. This is achieved thanks to a min-tree representation $\mathcal{TT}$ 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 τ, the lowest CC that contains τ and a smaller minimum is the lowest ancestor node whose smallest local minimum shape is different from τ.

To compute the final saliency map, we rely on the Khalimsky's grid, proposed by Khalimsky et al. (1990), and depicted in Fig. 6; it is composed of 0-faces (points), 1-faces (edges) and 2-faces (pixels). The saliency map $\mathcal{M}_\mathcal{E}$ is based on the extinction values, where we weigh the boundaries (0-faces and 1-faces) 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.

We refer the reader to Najman and Couprie (2006); Berger et al. (2007); Carlinet and Géraud (2014); Géraud et al. (2013) for details about the tree construction, and to Xu et al. (2015a) for details about the efficient computation of some information

```
COMPUTE_EXTINCTION_VALUE(𝒯, 𝒜)
(parent_𝒯, ℛ_𝒯) ← COMPUTE_TREE(𝒜);
for all τ ∈ 𝒯 do original_min(τ) ← undef;
for i ← 0 to N_𝒯 do
    τ ← R_𝒯(i), τ_p ← parent_𝒯(τ);
    if original_min(τ) = undef then original_min(τ) = τ;
    if original_min(τ_p) = undef then
        original_min(τ_p) ← original_min(τ);
    else
        if 𝒜(original_min(τ_p)) > 𝒜(original_min(τ)) then
            original_min(τ_p) ← original_min(τ);
for all τ ∈ 𝒯 do
    if τ is not a local minimum then
        ℰ(τ) ← 0;
    else
        τ_p ← parent_𝒯(τ);
        while original_min(τ_p) = τ & τ_p ≠ parent_𝒯(τ_p)
        do
            τ_p ← (parent_𝒯(τ_p));
        ℰ(τ) ← 𝒜(τ_p) − 𝒜(τ);
return 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 $\mathcal{TT}$. Note that *parent*$_\mathcal{T}$ encodes the min-tree $\mathcal{TT}$ constructed on the shape space.

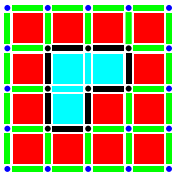


Fig. 6: Materialization of pixels with 0-faces (blue disks), 1-faces (green strips), and 2-faces (red squares). The original pixels are the 2-faces, the boundaries are materialized with 0-faces and 1-faces. The contour of the cyan region is composed of black 1-faces and 0-faces.

$\mathcal{A}$ (namely A, S_f, L, and S_∇).

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 Arbelaez et al. (2011). For the Weizmann segmentation database proposed by Alpert et al. (2012), 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 Alpert et al. (2012). Each image contains two objects to be segmented. We use the

COMPUTE_SALIENCY_MAP(f)
($parent, \mathcal{T}, \mathcal{A}$) ← COMPUTE_TREE(f);
$\mathcal{A}_{\lambda_s}$ ← COMPUTE_ATTRIBUTE($parent, \mathcal{T}, \mathcal{A}$);
$\lambda_s^M \leftarrow 0$;
for all $\tau \in \mathcal{T}$ **do** $\lambda_s^M \leftarrow \max(\lambda_s^M, \mathcal{A}_{\lambda_s}(\tau))$;
for all $\tau \in \mathcal{T}$ **do** $\mathcal{A}^{\downarrow}_{\lambda_s} \leftarrow \lambda_s^M - \mathcal{A}_{\lambda_s}(\tau)$;
$\mathcal{E} \leftarrow$ COMPUTE_EXTINCTION_VALUE($\mathcal{T}, \mathcal{A}^{\downarrow}_{\lambda_s}$);
for all 1-*face* e **do** $\mathcal{M}_{\mathcal{E}}(e) \leftarrow 0$;
for all e **do**
  $N_a \leftarrow appear(e), N_v \leftarrow vanish(e)$;
  while $N_a \neq N_v$ **do**
    $\mathcal{M}_{\mathcal{E}}(e) \leftarrow \mathbf{max}(\mathcal{E}(N_a), \mathcal{M}_{\mathcal{E}}(e)), N_a \leftarrow parent(N_a)$;
for all 0-*face* o **do**
  $\mathcal{M}_{\mathcal{E}}(o) \leftarrow \mathbf{max}(\mathcal{M}_{\mathcal{E}}(e_1), \mathcal{M}_{\mathcal{E}}(e_2), \mathcal{M}_{\mathcal{E}}(e_3), \mathcal{M}_{\mathcal{E}}(e_4))$;
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.

color tree of shapes proposed by Carlinet and Géraud (2015), 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 Monasse and Guichard (2000) 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 $\mathcal{A}^{\downarrow}_{\lambda_s}$ are depicted on the bottom of this figure. They represent hierarchical pre-segmentations.

Some illustrations of the extinction-based saliency map $\mathcal{M}_{\mathcal{E}}$ applied on images in the dataset of BSDS500 Arbelaez et al. (2011) 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, salient level lines are highlighted in $\mathcal{M}_{\mathcal{E}}$ employing the attribute $\mathcal{A}^{\downarrow}_{\lambda_s}$. 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.hierarchymsll`.

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

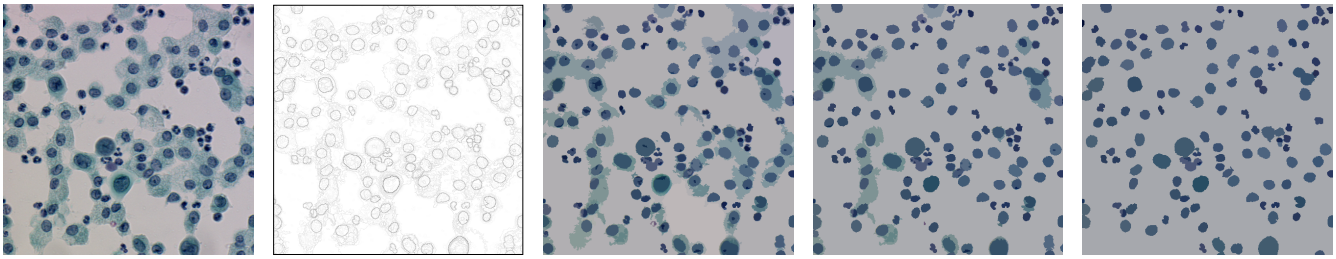
We have also evaluated our hierarchical image simplifications in context of segmentation on Weizmann segmentation evaluation database in Alpert et al. (2012). For the 100 images containing 2 objects in this database (See Fig. 7 for several examples), the saliency maps are thresholded with a fixed thresholding 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

Fig. 7: Some pre-segmentation results obtained with our proposed method on the segmentation evaluation database in Alpert et al. (2012). Top: input images; Middle: pre-segmentations obtained with the simplification method; Bottom: inverted saliency maps for hierarchical simplifications.

computed by Géraud et al. (2013) on grayscale versions of the input images f. We performed two tests as presented in Alpert et al. (2012) 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 Alpert et al. (2012) 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 Dollár and Zitnick (2015) (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 explicitely 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 similar. However, our method, as a pre-segmentation method without using any texture information has



(a) Input image. (b) Saliency map. (c) Slightly simplified. (d) Moderately simplified. (e) Strongly simplified.

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

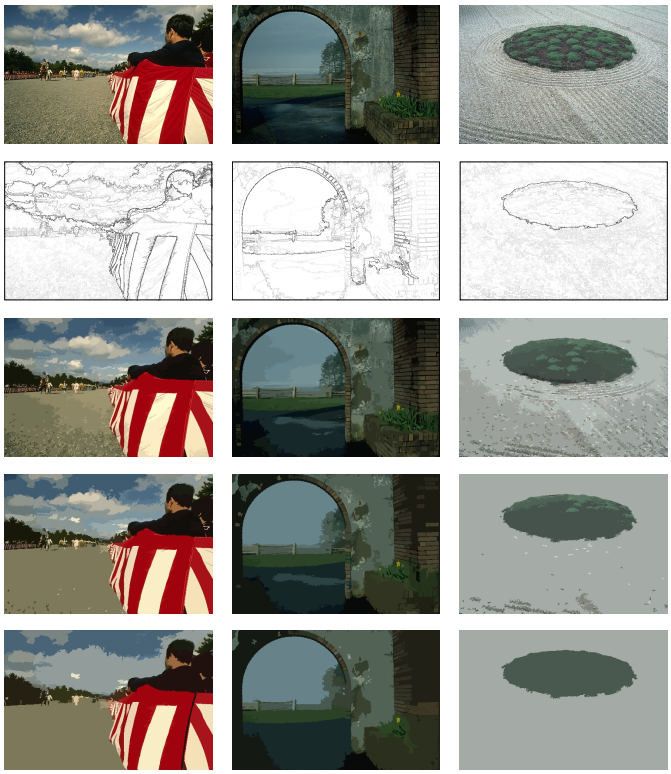


Fig. 8: Illustration of the hierarchical image simplification using the attribute function $\mathcal{A}^{\downarrow}_{\lambda_s}$, applied on some images from the dataset of BSDS500 proposed by Arbelaez et al. (2011). From top to bottom: input images; inverted saliency maps; slight simplification; moderate simplification; strong simplification.

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.hierarchymsll`.

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 Salembier and Wilkinson (2009). In all these strategies, the attribute function $\mathcal{A}$ characterizing each node

Table 1: Results of single segment coverage test using F-measure on the two objects dataset in Alpert et al. (2012). "Our2" stands for our method with the attribute $\mathcal{A}_\nabla$ using the gradient computed by Dollár and Zitnick (2015).

Method	Average	Larger	Smaller
Our2	***0.80***	***0.79***	***0.81***
Gpb-owt-ucm	**0.79**	**0.80**	**0.78**
Our	*0.77*	*0.77*	*0.76*
Gpb-owt-ucm without texture	0.76	0.76	0.75
Gpb in Alpert et al. (2012)	0.72	0.70	0.75
Method in Alpert et al. (2012)	0.68	0.70	0.65
SWA in Sharon et al. (2006)	0.66	0.74	0.57
MeanShift	0.61	0.65	0.58
N-Cuts	0.58	0.66	0.49

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.*, Xu et al. (2015b)). 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 Guigues et al. (2006), 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 λ_s^+ and λ_s^- for each region R of the input hierarchy H, where λ_s^+ (*resp.*, λ_s^-) corresponds to the smallest parameter λ_s such that the region $R \in H$ belongs (*resp.*. does not belong) to the optimal solution of segmentation by minimizing E_{λ_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_{λ_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

Table 2: Fragmented coverage test results for the two objects dataset proposed by Alpert et al. (2012): 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 Dollár and Zitnick (2015).

Method	Averaged		Larger object		Smaller object	
	F-measure	#fragments	F-measure	#fragments	F-measure	#fragments
SWA in Sharon et al. (2006)	**0.88**	3.13	**0.91**	3.88	**0.84**	2.37
Our2	***0.86***	*2.40*	*0.85*	*3.00*	***0.86***	*1.81*
Method in Alpert et al. (2012)	0.85	**1.67**	0.87	**2.00**	**0.84**	**1.33**
N-Cuts in Shi and Malik (2000)	0.84	2.64	**0.88**	3.34	0.80	1.93
Gpb reported in Alpert et al. (2012)	0.84	2.95	0.87	3.60	0.81	2.30
Our	*0.83*	*3.16*	*0.85*	*4.10*	*0.81*	*2.23*
Gpb-owt-ucm in Arbelaez et al. (2011)	0.82	**1.57**	0.84	**1.79**	0.81	**1.35**
Gpb-owt-ucm without texture	0.81	2.72	0.82	3.32	0.80	2.12
MeanShift in Comaniciu and Meer (2002)	0.78	3.65	0.85	4.49	0.71	2.81

and extended by Kiran and Serra (2014). These methods work on an input hierarchy of segmentations, which is very different from the tree of shapes (a natural and equivalent image representation). Indeed, each cut of the tree of shapes is a subset of the image domain, while each cut of the hierarchy of segmentations forms a partition of the image domain; see Ronse (2014). This basic difference prohibits the direct use on the tree of shapes of the classical works which find optimal hierarchical segmentations by energy minimization. In this sense, our approach can be seen as an extension of the scale-set theory proposed in Guigues et al. (2006) to the tree of shapes.

Another related work is the one in Ballester et al. (2007). It also selects meaningful level lines for image simplification and segmentation 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 Ballester et al. (2007) 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 Ballester et al. (2007) is computationally expensive. We propose to fix that issue thanks to a reasonable ordering of level lines based on their quantitative meaningfulness measurement (*e.g.*, the average of gradient's magnitude along the level line $\mathcal{A}_\nabla$). 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 Ballester et al. (2007) 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 Ballester et al. (2007). Our approach is almost linear w.r.t. the number of nodes in the tree. Yet, the method of Ballester et al. (2007) seems to depend also on the depth of the tree. In Ballester et al. (2007), the authors proposed applying the simplification scheme successively with a set of augmenting parameters λ_s so that to construct the input hierarchy. Then they employed the scheme of scale-set theory proposed by Guigues et al. (2006) on the obtained hierarchy to achieve a final hierarchy of optimal segmentations. In our case, rather than using a fixed parameter λ_s or a set of fixed parameters, we propose to assign a measure related to λ_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.

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

Image	Depth	#Nodes	Time (s)	
			Ballester et al. (2007)	Our
House	126	23588	4.11	0.22
Camera	126	24150	4.19	0.23
Lena	141	84699	27.77	0.92
Peppers	176	97934	48.18	0.93
Boat	255	100518	87.24	0.94
Barbara	131	106285	51.87	0.99
Mandrill	185	153029	200.22	1.34

It is worth noticing that the minimization of Mumford-Shah-like functional has also been applied to shape analysis in Tari and Genctav (2014). 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 Tari and Genctav (2014), 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, 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 $\mathcal{A}_{\lambda_s}$ 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.hierarchymsll`.

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 appicable to our case, probability map incorporating region features, such as proposed in Bai and Sapiro (2009), are worth exploring. Computing the tree of shapes of such probability maps has already been proved valuable (see Dubrovina et al. (2014)). 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 Salembier and Garrido (2000); Ballester et al. (2007)) and the energy based on spectral unmixing used for hyperspectral image segmentation in Veganzones et al. (2014). 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

Alpert, S., Galun, M., Brandt, A., Basri, R., 2012. Image segmentation by probabilistic bottom-up aggregation and cue integration. IEEE Transactions on Pattern Analysis and Machine Intelligence 34, 315–327.

Arbelaez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 33, 898–916.

Bai, X., Sapiro, G., 2009. Geodesic matting: A framework for fast interactive image and video segmentation and matting. International Journal of Computer Vision 82, 113–132.

Ballester, C., Caselles, V., Igual, L., Garrido, L., 2007. Level lines selection with variational models for segmentation and encoding. Journal of Mathematical Imaging and Vision 27, 5–27.

Berger, C., Géraud, T., Levillain, R., Widynski, N., Baillard, A., Bertin, E., 2007. Effective component tree computation with application to pattern recognition in astronomical imaging, in: Proc. of IEEE International Conference on Image Processing, pp. 41–44.

Cao, F., Musé, P., Sur, F., 2005. Extracting meaningful curves from images. Journal of Mathematical Imaging and Vision 22, 159–181.

Cardelino, J., Randall, G., Bertalmio, M., Caselles, V., 2006. Region based segmentation using the tree of shapes, in: Proc. of IEEE International Conference on Image Processing, pp. 2421–2424.

Carlinet, E., Géraud, T., 2014. A comparative review of component tree computation algorithms. IEEE Transactions on Image Processing 23, 3885–3895.

Carlinet, E., Géraud, T., 2015. MToS: A tree of shapes for multivariate images. IEEE Transactions on Image Processing 24, 5330–5342.

Caselles, V., Coll, B., Morel, J.M., 1999. Topographic maps and local contrast changes in natural images. International Journal of Computer Vision 33, 5–27.

Comaniciu, D., Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 603–619.

Dollár, P., Zitnick, C.L., 2015. Fast edge detection using structured forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 37, 1558–1570.

Dubrovina, A., Hershkovitz, R., Kimmel, R., 2014. Image editing using level set trees, in: Proc. of IEEE International Conference on Image Processing, pp. 4442–4446.

Felzenszwalb, P., Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. International Journal of Computer Vision 59, 167–181.

Géraud, T., Carlinet, E., Crozet, S., Najman, L., 2013. A quasi-linear algorithm to compute the tree of shapes of nD images, in: Proc. of International Symposium on Mathematical Morphology, Springer. pp. 98–110.

Guigues, L., Cocquerez, J.P., Le Men, H., 2006. Scale-sets image analysis. International Journal of Computer Vision 68, 289–317.

Khalimsky, E., Kopperman, R., R Meyer, P., 1990. Computer graphics and connected topologies on finite ordered sets. Topology and its Applications 36, 1–17.

Kiran, B.R., Serra, J., 2014. Global-local optimizations by hierarchical cuts and climbing energies. Pattern Recognition 47, 12–24.

Koenderink, J.J., 1984. The structure of images. Biological Cybernetics 50, 363–370.

Lu, H., Woods, J.C., Ghanbari, M., 2007. Binary partition tree analysis based on region evolution and its application to tree simplification. IEEE Transactions on Image Processing 16, 1131–1138.

Monasse, P., Guichard, F., 2000. Fast computation of a contrast-invariant image representation. IEEE Transactions on Image Processing 9, 860–872.

Mumford, D., Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics 42, 577–685.

Najman, L., 2011. On the equivalence between hierarchical segmentations and ultrametric watersheds. Journal of Mathematical Imaging and Vision 40, 231–247.

Najman, L., Couprie, M., 2006. Building the component tree in quasi-linear time. IEEE Transactions on Image Processing 15, 3531–3539.

Najman, L., Schmitt, M., 1996. Geodesic saliency of watershed contours and hierarchical segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 1163–1173.

Pardo, A., 2002. Semantic image segmentation using morphological tools, in: Proc. of IEEE International Conference on Image Processing, pp. 745–748.

Perret, B., Cousty, J., Ura, J.C.R., Guimarães, S.J.F., 2015. Evaluation of morphological hierarchies for supervised segmentation, in: Proc. of International Symposium on Mathematical Morphology, Springer. pp. 39–50.

Ronse, C., 2014. Ordering partial partitions for image segmentation and filtering: Merging, creating and inflating blocks. Journal of Mathematical Imaging and Vision 49, 202–233.

Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation and information retrieval. IEEE Transactions on Image Processing 9, 561–576.

Salembier, P., Wilkinson, M.H., 2009. Connected operators. Signal Processing Magazine 26, 136–157.

Sharon, E., Galun, M., Sharon, D., Basri, R., Brandt, A., 2006. Hierarchy and adaptivity in segmenting visual scenes. Nature 442, 810–813.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 888–905.

Tari, S., Genctav, M., 2014. From a non-local ambrosio-tortorelli phase field to a randomized part hierarchy tree. Journal of mathematical imaging and vision 49, 69–86.

Vachier, C., Meyer, F., 1995. Extinction value: a new measurement of persistence, in: Proc. of IEEE Workshop on Nonlinear Signal and Image Processing, pp. 254–257.

Veganzones, M.A., Tochon, G., Dalla-Mura, M., Plaza, A.J., Chanussot, J., 2014. Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation. IEEE Transactions on Image Processing 23, 3574–3589.

Xu, Y., Carlinet, E., Géraud, T., Najman, L., 2015a. Efficient computation of attributes and saliency maps on tree-based image representations, in: Proc. of International Symposium on Mathematical Morphology, Springer. pp. 693–704.

Xu, Y., Géraud, T., Najman, L., 2013a. Salient level lines selection using the mumford-shah functional, in: Proc. of IEEE International Conference on Image Processing, pp. 1227–1231.

Xu, Y., Géraud, T., Najman, L., 2013b. Two applications of shape-based morphology: Blood vessels segmentation and a generalization of constrained connectivity, in: Proc. of International Symposium on Mathematical Morphology, Springer. pp. 390–401.

Xu, Y., Géraud, T., Najman, L., 2015b. Connected filtering on tree-based shape-spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence To be published.