AN EFFICIENT HIERARCHICAL GRAPH BASED IMAGE SEGMENTATION

Silvio Jamil F. Guimarães[‡§]   *Jean Cousty*[§]*, Yukiko Kenmochi*[§] *and Laurent Najman*[§]

[‡]VIPLAB - ICEI - PUC Minas
sjamil@pucminas.br

[§]Université Paris-Est, LIGM, ESIEE - UPEMLV - CNRS
{j.cousty,y.kenmochi,l.najman} @esiee.fr

ABSTRACT

Hierarchical image segmentation provides region-oriented scale-space, *i.e.*, a set of image segmentations at different detail levels in which the segmentations at finer levels are nested with respect to those at coarser levels. Most image segmentation algorithms, such as region merging algorithms, rely on a criterion for merging that does not lead to a hierarchy, and for which the tuning of the parameters can be difficult. In this work, we propose a hierarchical graph based image segmentation relying on a criterion popularized by Felzenzwalb and Huttenlocher. We illustrate with both real and synthetic images, showing efficiency, ease of use, and robustness of our method.

Index Terms— Hierarchical image segmentation, Edge-weighted graph, Saliency map

1. INTRODUCTION

Image segmentation is the process of grouping perceptually similar pixels into regions. A hierarchical image segmentation is a set of image segmentations at different detail levels in which the segmentations at coarser detail levels can be produced from simple merges of regions from segmentations at finer detail levels. Therefore, the segmentations at finer levels are nested with respect to those at coarser levels. Hierarchical methods have the interesting property of preserving spatial and neighboring information among segmented regions. Here, we propose a hierarchical image segmentation in the framework of edge-weighted graphs, where the image is equipped with an adjacency graph and the cost of an edge is given by a dissimilarity between two points of the image.

Any hierarchy can be represented with a minimum spanning tree. The first appearance of this tree in pattern recognition dates back to the seminal work of Zahn [1]. Lately, its use for image segmentation was introduced by Morris *et al.* [2] in 1986 and popularized in 2004 by Felzenswalb and Huttenlocher [3]. However the region-merging method [3] does not provide a hierarchy. In [4, 5], it was studied some optimality properties of hierarchical segmentations. Considering that, for a given image, one can tune the parameters of the well-known method [3] for obtaining a correct segmentation of this image. We provide in this paper a hierachical version of this method that removes the need for parameter tuning.

The algorithm of [3] is the following. First, a minimum spanning tree (MST) is computed, and all the decisions are taken on this tree. For each edge linking two vertices x and y, following a non-decreasing order of their weights, the following steps are performed:

(i) Find the region X that contains x.
(ii) Find the region Y that contains y.
(iii) Merge X and Y according to a certain criterion.

The criterion for region-merging in [3] measures the evidence for a boundary between two regions by comparing two quantities: one based on intensity differences across the boundary, and the other based on intensity differences between neighboring pixel within each region. More precisely, in step (iii), in order to know whether two regions must be merged, two measures are considered. The *internal difference* $Int(X)$ of a region X is the highest weight of an edge linking two vertices of X in the MST. The *difference* $Diff(X, Y)$ between two neighboring regions X and Y is the smallest weight of an edge that links X to Y. Then, two regions X and Y are merged when:

$$Diff(X, Y) \leq \min\{Int(X) + \frac{k}{|X|}, Int(Y) + \frac{k}{|Y|}\} \qquad (1)$$

where k is a parameter allowing to prevent the merging of large regions (*i.e.*, larger k force smaller regions to be merged).

The merging criterion defined by Eq. (1) depends on the scale k at which the regions X and Y are observed. More precisely, let us consider the *(observation) scale* $S_Y(X)$ of X relative to Y as a measure based on the difference between X and Y, on the internal difference of X and on the size $|X|$ of X:

$$S_Y(X) = (Diff(X, Y) - Int(X)) \times |X|. \qquad (2)$$

Then, the scale $S(X, Y)$ is simply defined as:

$$S(X, Y) = \max(S_Y(X), S_X(Y)). \qquad (3)$$

Thanks to this notion of a scale Eq. (1) can be written as:

$$k \geq S(X, Y). \qquad (4)$$

In other words, Eq.(4) states that the neighboring regions X and Y merge when their scale is less than the threshold parameter k.

Even if the image segmentation results obtained by the method proposed in [3] are interesting, the user faces two major issues:

- first, the number of regions may increase when the parameter k increases. This should not be possible if k was a true scale of observation: indeed, it violates the *causality principle* of multi-scale analysis, that states in our case [6] that a contour present at a scale k_1 should be present at any scale $k_2 < k_1$. Such a behaviour is demonstrated on Fig. 1.

- Second, even when the number of regions decreases, contours are not stable: they can move when the parameter k varies, violating a *location principle*. Such a situation is illustrated on Fig. 2.

Given these two issues, the tunning of the parameters of [3] is a difficult task.

Following [6], we believe that, in order for k to be a true scale-parameter, we have to satisfy both the causality principle and the location principle, which leads to work with hierarchy of segmentations. Reference [7] is the first to propose an algorithm producing a hierarchy of segmentations based on [3]. However, this method is an
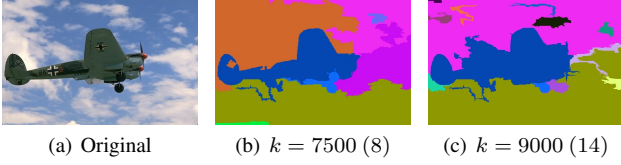
(a) Original   (b) $k = 7500$ (8)   (c) $k = 9000$ (14)

Fig. 1. A real example illustrating the violation of the causality principle by [3]: the number of regions (in parentheses) increases from 8 to 14, instead of decreasing when the so-called "scale of observation" increases.
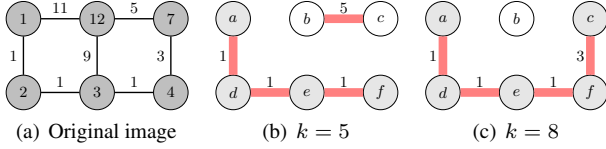

(a) Original image   (b) $k = 5$ (c) $k = 8$

Fig. 2. An example illustrating the violation of the location property by [3]: the contours are unstable from one "scale" to another.

iterative version of [3] that uses a threshold function, and requires a tunning of the threshold parameter.

The main result of this paper is an efficient hierarchical image segmentation algorithm based on the dissimilarity measure of [3]. Our algorithm has a computational cost similar to [3], but provides all scales of observations instead of only one segmentation level. As it is a hierarchy, the result of our algorithm satisfies both the locality principle and the causality principle. In particular, and in contrast with [3], the number of regions is decreasing when the scale parameter increases, and the contours do not move from one scale to another.

Figure 3 illustrates the results obtained by applying our method to the same image of Fig. 1(a), with segmentations at two different scales of observations, as well as a saliency map [8, 4, 5] (a map indicating the disparition level of contours and whose thresholds give the set of all segmentations).
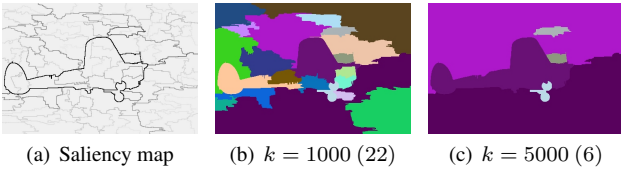

(a) Saliency map   (b) $k = 1000$ (22)   (c) $k = 5000$ (6)

.

Fig. 3. A real example illustrating the saliency map of Fig. 1(a) computed with our approach. We display in (b) and (c) two image segmentations extracted from the hierarchy at scales 1000 and 5000, together with their number of regions (in parentheses).

This work is organized as follows. In Section 2, we present our hierarchical method for color image segmentation. Some experimental results are given in Section 3. Finally, in Section 4, some conclusions are drawn and further works are discussed.

2. AN EFFICIENT HIERARCHICAL GRAPH BASED IMAGE SEGMENTATION

In this section, we describe our method to compute a hierarchy of partitions based on observation scales as defined by Eq. 3. Let us first recall some important notions for handling hierarchies [2, 4, 5].

To every tree T spanning the set V of the image pixels, to every map $w : E \to \mathbb{N}$ that weights the edges of T and to every threshold $\lambda \in \mathbb{N}$, one may associate the partition $\mathcal{P}_\lambda^w$ of V induced by the connected components of the graph made by V and the edges of weight below λ. It is well known [2, 5] that for any two values λ_1 and λ_2 such that $\lambda_1 \geq \lambda_2$, the partitions $\mathcal{P}_{\lambda_1}^w$ and $\mathcal{P}_{\lambda_2}^w$ are *nested* and $\mathcal{P}_{\lambda_1}^w$ is *coarser* than $\mathcal{P}_{\lambda_2}^w$. Hence, the set $\mathcal{H}^w = \langle \mathcal{P}_\lambda^w \mid \lambda \in \mathbb{N} \rangle$ is a *hierarchy of partitions induced by the weight map w.*

Our algorithm does not explicitly produce a hierarchy of partitions, but instead it produces a weight map L (scales of observations) from which the desired hierarchy $\mathcal{H}^L$ can be infered. It starts from a minimum spanning tree T of the edge-weighted graph built from the image. In order to compute the scale $L(e)$ associated with each edge of T, our method iteratively considers the edges of T in a non-decreasing order of their weights. For every edge e, the weight map $L(e)$ is initialized to ∞; then for each edge e linking two vertices x and y the following steps are performed:

  (i) Find the the region X of $\mathcal{P}_{w(e)}^w$ that contains x.

  (ii) Find the the region Y of $\mathcal{P}_{w(e)}^w$ that contains y.

  (iii) Compute the hierarchical observation scale $L(e)$.

At step (iii), the *hierarchical scale* $S'_Y(X)$ of X relative to Y is needed to obtain the value $L(e)$. Intuitively, $S'_Y(X)$ is the lowest observation scale at which some sub-region of X, namely X^*, will be merged to Y. More precisely, using an internal parameter v, this scale is computed as follows:

  (1) Initialize the value of v to 0.

  (2) Increment the value of v by 1.

  (3) Find the the region X^* of $\mathcal{P}_v^L$ that contains x.

  (4) Repeat steps 2 and 3 while $S_Y(X^*) > v$

  (5) $S'_Y(X) = v$.

With the appropriate changes, the same algorithm allows $S'_X(Y)$ to be computed. Then, the hierarchical scale $L(e)$ is simply set to:

$$L(e) = \max\{S'_Y(X), S'_X(Y)\}. \tag{5}$$

Fig. 4 illustrates the result of our method on a pedagogical example. Starting from the graph of Fig. 4(a), our method produces the hierarchical observation scales depicted in Fig. 4(b). As for the method of [3], our algorithm only considers the edges of the minimum spanning tree (see Fig. 4(c)). The whole hierarchy is depicted as a dendrogram in Fig. 4(d), whereas two levels of the hierarchy (at scales 2 and 9) are shown in Fig. 4(e) and (f).

Let us illustrate the computation of a hierarchical obervation scale on the graph of Fig. 5(a). To this end, we consider the iteration of the algorithm at which the edge e linking B to G is analyzed. At this step, the edges of the MST of weight below $w(e) = 10$ have been already processed. Therefore, the hierarchical observation scale of these edges (depicted by continuous lines in the figure) is already known as shown in 5(b). The regions X and Y obtained at steps (i) and (ii) are set to $\{A, B, C, D, E\}$ and $\{F, G, H, I\}$ respectively. Then, in order to find the value $L(e)$ at step (iii), the partitions $\mathcal{P}_i^w$ for $i = \{2, 7, 13, 18\}$ must be considered. We have: $\mathcal{P}_2^L = \{\{B, C\}, \{A, D\}, \{E\}, \{G\}, \{F\}, \{I, H\}\}, \mathcal{P}_7^L = \{\{B, C\},$
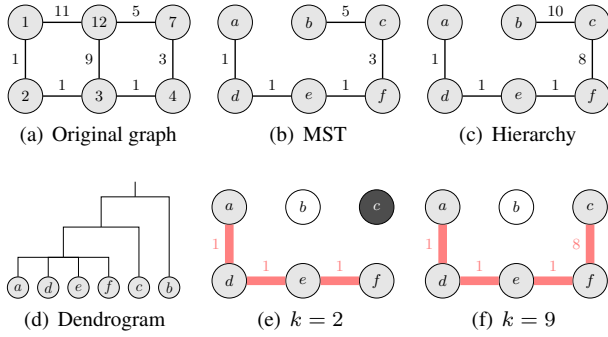
(a) Original graph   (b) MST   (c) Hierarchy

(d) Dendrogram   (e) $k = 2$ (f) $k = 9$

Fig. 4. Example of hierarchical image segmentations. In contrast to example in Fig. 2, the contours are stable from a scale to another, providing a hierarchy.
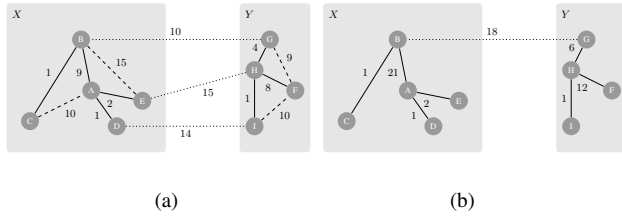

(a)   (b)

Fig. 5. Example for computing the hierarchical scale for an edge-weighted graph. For this example, we suppose that all scales for the regions X and Y are already computed, and we will calculate the hierarchical scale for the edge connectinge B and G.

$\{A, D, E\}, \{F\}, \{G, H, I\}\}$, $\mathcal{P}_{13}^L = \{\{B, C\}, \{A, D, E\}, \{F, G, H, I\}\}$ and $\mathcal{P}_{18}^L = \{\{B, C\}, \{A, D, E\}, \{F, G, H, I\}\}$. By the application of steps (1-5), the value $S'_Y(X)$ is found to be 18 since 18 is the first value below the observation scale of the region containing B relatively to Y. The same process is made for $S'_X(Y)$, but the regions are $\{G\}$, $\{G, H, I\}$ and $\{G, H, I, F\}$. Moreover, the observation scale is 12 since 12 is the first value below the observation scale of the region containing G relatively to X. Finally, the observation scale of X and Y is 18.

To efficiently implement our method, we use some data structures similar to the ones proposed in [5]; in particular, the management of the collection of partitions are due to Tarjan's union find. Furthermore, we made some algorithmic optimizations to speed up the computations of the observation scales. In order to illustrate an example of computation time, we implemented all our algorithm in C++ on a standard single CPU computer under windows Vista, we run it in a Intel Core 2 Duo, 4GB. For the image illustrated in Fig. 1(a) (with size 321x481), the hierarchy is computed in 2.7 seconds, and the method proposed in [3] spent 1.3 seconds .

3. EXPERIMENTAL RESULTS

A major difficulty of experiments is the design of an adequate edge-cost, well adapted to the content to be segmented. A practical solution is to use some dissimilarity functions, and many different approaches are used in the litterature. In this work, the underlying
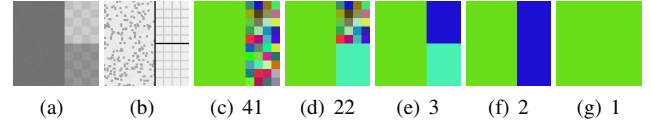

(a)   (b)   (c) 41   (d) 22   (e) 3   (f) 2   (g) 1

Fig. 6. An example of a hierarchical image segmentations of a synthetic image containinng three perceptually big regions. The saliency map of the image (a) is showned in (b). The number of regions of the segmented images is written under each figure.
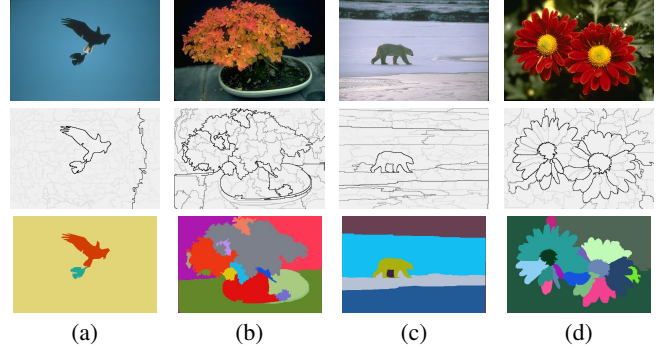

(a)   (b)   (c)   (d)

Fig. 7. Top row: some images of the Berkeley database [9]. Middle row: saliency maps of these images. The numbers of scales of these hierarchies are (a) 240, (b) 429, (c) 405 and (d) 443.Bottom row: according to our subjective judgment, the best segmentations extracted from the hierarchies. The numbers of regions are (a) 3, (b) 16, (c) 6 and (d) 18.

graph is the one induced by the 4-adjacency relation, and the edges are weighted by a simple color gradient computed by an Euclidean distance in the RGB space.

In Fig. 6, we present some results on an artifical image containing three perceptually big regions. On this example, one can easily verify the hierarchical property of our method by looking at the segmentations at scales *resp.* 1000, 2000, 5000, 140000 and 224000 (*resp.* Fig. 6(c), (d), (e), (f) and (g)). Since the resulting segmentations are nested, the whole hierarchy can be presented in a saliency map (see Fig. 6(b)).

Fig. 7 illustrates the performance of our method on some images of the Berkeley's database [9]. Note that, as in [3], an area filtering is applied to eliminate small regions (smaller than 500 pixels).

Comparison of the results of our algorithm with the ones of [3] are difficult, since the tuning of the parameters of [3] is critical and since we produce a whole hierarchy of segmentations. We made three experiments. First, we try to set the correct parameter for [3], *i.e.* the parameter that produces the best (subjective) visual result (Fig. 8). We can compare this result with on the one hand, the "best" segmentation extracted from our hierarchy in Fig. 7, and on the other hand, with a segmentation from our hierarchy containing the same number of regions as [3] (Fig. 8). In a second experiments, we fixed the number of regions to 15 for all images, and tune the parameter of [3] to obtain this number of regions. We can compare these segmentations with our own results on Fig. 9. The last experiments is designed to assess the robustness to random impulse noise, see Fig. 10.
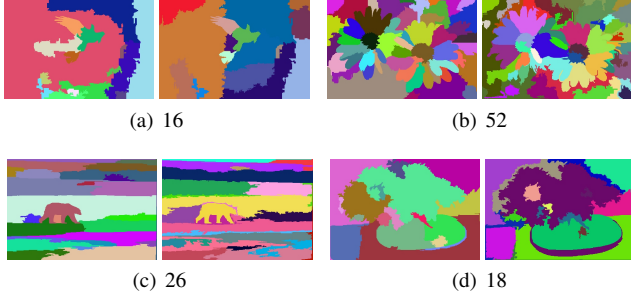
(a) 16   (b) 52

(c) 26   (d) 18

Fig. 8. Comparison between [3] and our approach. For each pair of images, the right image shows the best result (according to our judgment and our experiments) from [3] and the left image shows a segmentation extracted from our hierarchical result, with the same number of regions.
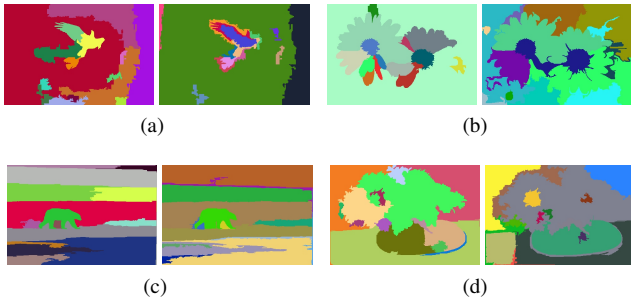

(a)   (b)

(c)   (d)

Fig. 9. Examples of image segmentation where the number of regions has been set to 15. For each pair of images, the left one shows a segmentation extracted from our hierarchy, with the desired number of regions; and the right one shows the result obtained with [3] by varying the parameter k until the desired number of regions is found.
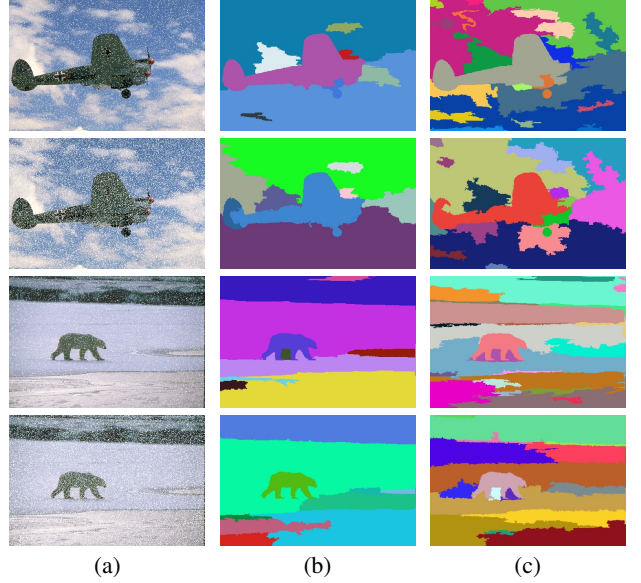

(a)   (b)   (c)

Fig. 10. Examples of segmentations for images corrupted by a random salt noise. The corrupted images (at different levels - 70% and 90%) are shown on the first column. The results of our method and [3] are illustrated in the second and third columns, respectively.

4. CONCLUSIONS

This paper proposes an efficient hierarchical segmentation method based on the observation scales of [3]. In contrast to [3], our method produces the complete set of the segmentations at every scales, and satisfies both the causality and location principle defined by [6]. An important practical consequence of these properties is to ease the selection of a scale level adapted to a particular task. We visually assessed our method on some real images by comparing our segmentations to those of [3]. Even if more (quantitative) tests (such as the ones proposed by [10]) are needed for drawing definitive conclusions, the produced segmentations are promizing, in particular w.r.t. robustness. As future work, we will investigate using more information into the definition of observation scale as well as learning which information is pertinent for a given practical task.

5. REFERENCES

[1] C. T. Zahn, "Graph-theoretical methods for detecting and describing gestalt clusters," *IEEE Trans. Comput.*, vol. 20, pp. 68–86, January 1971.

[2] O.J. Morris, M.de J. Lee, and A.G. Constantinides, "Graph theory for image analysis: an approach based on the shortest spanning tree," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 133, no. 2, pp. 146 –152, april 1986.

[3] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. 59, pp. 167–181, September 2004.

[4] Laurent Najman, "On the equivalence between hierarchical segmentations and ultrametric watersheds," *JMIV*, vol. 40, pp. 231–247, 2011.

[5] Jean Cousty and Laurent Najman, "Incremental algorithm for hierarchical minimum spanning forests and saliency of watershed cuts," in *ISMM*, vol. 6671 of *LNCS*, pp. 272–283. Springer, 2011.

[6] Laurent Guigues, Jean Pierre Cocquerez, and Hervé Le Men, "Scale-sets image analysis," *IJCV*, vol. 68, no. 3, pp. 289–317, 2006.

[7] Yll Haxhimusa and Walter Kropatsch, "Segmentation graph hierarchies," in *Structural, Syntactic, and Statistical Pattern Recognition*, vol. 3138 of *LNCS*, pp. 343–351. Springer, 2004.

[8] L. Najman and M. Schmitt, "Geodesic saliency of watershed contours and hierarchical segmentation," *PAMI*, vol. 18, no. 12, pp. 1163–1173, December 1996.

[9] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV*, July 2001, vol. 2, pp. 416–423.

[10] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik, "Contour detection and hierarchical image segmentation," *PAMI*, vol. 33, pp. 898–916, 2011.