14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

MULTI-SCALE IMAGE SEGMENTATION IN A HIERARCHY OF PARTITIONS

O. Lezoray, C. Meurie, P. Belhomme and A. Elmoataz

LUSAC EA 2607, Vision and Image Analysis, IUT SRC, 120 Rue de l'exode, F-50000 SAINT-LÔ, FRANCE
phone: + (33) 0233775517, fax: + (33)0233771167, email: Olivier.Lezoray@info.unicaen.fr
web: http://www.stlo.unicaen.fr/vai

ABSTRACT

In this paper, we propose a new multi-scale image segmentation relying on a hierarchy of partitions. First of all, we review morphological methods based on connections which produce hierarchical image segmentations and introduce a new way of generating fine segmentations. From connection-based fine partitions, we focus on producing hierarchical segmentations. Multi-scale segmentations obtained by scale-space or region merging approaches have both benefits and drawbacks, therefore we propose to integrate a scale-space approach in the production of a hierarchy of partitions which merges similar regions. Starting from an initial over-segmented fine partition, a region adjacency graph is alternatively simplified and decimated. The simplification of graph nodes models tends to produce similar ones and this is used to merge them. The algorithm can be used to simplify the image at low scales and to segment it at high scales.

1. INTRODUCTION

Image segmentation consists in partitioning an image in more or less regular or coherent zones according to a given criterion. Thus one usually seeks an image partition in zones the values of which follow a given model of organization. Automatic segmentation of images is a central problem in image analysis since a partition of the image in regions makes the extraction of the primary visual components from an image possible, the latter being used to identify and recognize objects of interest. However there is a gap between the image itself and its description as a sole partition of the image into several regions. A way of circumventing this gap is to concentrate on region segmentation on the one hand and on the other hand on perceptual groupings extracted by a hierarchical vision of images. The union of regions is a group of elements which also is a region at a given scale, with local inner properties and global ones according to its neighborhood. Most of the time, low-level image segmentation algorithms cannot directly cope with this semantic gap since it is very difficult to directly construct the best image partition (if there is one). Thus it is necessary to deal with hierarchical methods which produce a multi-scale image segmentation. There are many ways of producing hierarchical partitions of an image and we propose to integrate in a single approach two of them which are usually not used together: scale-space and region merging. Multi-scale segmentations obtained by scale-space or region merging approaches have both benefits and drawbacks, therefore we propose to integrate a scale-space approach in the production of a hierarchy of partitions which merges similar regions. Starting from an initial over-segmented fine partition, a region adjacency graph is alternatively simplified and decimated. The simplification of graph nodes models tends to produce similar ones and this is used to merge them. The algorithm can be used to simplify the image at low scales and to segment it at high scales. In the next section we focus on how to obtain fine partitions and propose a new way of generating such ones. Third section presents the proposed multi-scale image segmentation and fourth section concludes.

2. HIERARCHY OF PARTITIONS

The traditional problem of the automatic segmentation of images is generally considered like a division of the image in disjoined areas, the result being a partition of the image domain. An image I is a set of pixels $I = \{p_1, p_2, ..., p_n\}$ and a region R is a subset of the image pixels composed of $|R|$ pixels.

Definition 1 *A partition P is a set of regions $P = \{R_1, R_2, ..., R_k\}$ so that: (1) the union of the partition regions provides the initial set: $I = \bigcup_{i=1}^{k} R_i$, (2) regions are disjoined: $\forall i, j, i \neq j, R_i \cap R_j = \emptyset$.*

Then one can define an ordering relation between two partitions: a partition P is included in a partition Q if every region R_j^P is completely included in a region R_i^Q. This defines a hierarchy of nested partitions of an image. Let H be a set of partitions associated to an image, H is a hierarchy of partitions if it is possible to define an inclusion order between any pair of H elements.

Definition 2 *A hierarchy of (nested) partitions is a set of partitions $H = \{P_1, P_2, ..., P_l\}$ so that the regions of the partition $P_i = \{R_1^i, R_2^i, ..., R_k^i\}$ are all included in the regions of the partition $P_j = \{R_1^j, R_2^j, ..., R_{k'}^j\}$ with $j > i, k > k'$ and $R_m^i \subseteq R_p^j$ or $R_m^i \cap R_p^j = \emptyset$.*

It means that two regions from two different partitions of a hierarchy are either disjoined or included one in the other. P_i is called the i^{th} level of the hierarchy, P_0 if the lowest one and is the finest partition, P_l is the highest level of the hierarchy and is the coarsest partition. The regions of the lowest level being always included in higher level regions, the regions of the $(i+1)^{th}$ partition can be obtained by merging ones of the i^{th} partition. Therefore, a hierarchy of partitions is naturally represented by a stack of Region Adjacency Graphs (RAGs) also called an irregular pyramid [1]. Links between regions that merge from one level to the next one are contained in a so-called contraction kernel [2].

In mathematical morphology, to have an ordering relation between successive levels of the hierarchy implies that the latter forms a complete lattice. The main morphological criteria that define hierarchies of partitions are based on connections (connective criteria) [3]. This enables to divide an image into zones according to a given criterion. For instance, an image I is divided into flat and connected zones when a partition P_i is created, so that for every x, the region $R \in P_i$ with $x \in R$ is the highest connected component which includes x and where the image I is constant and always equal to x. Whatever the connection criterion there always is a way of partitioning an image into regions that fulfill that criterion. The main morphological connective criteria are the flat zones and the watershed. The flat zones of an image I are the maximal connected components having a constant value, MEYER further introduced the quasi flat zones principle, a threshold connection [4].

Definition 3 *Two points p and q belong to the same quasi flat zone of an image I if there is a connected path $(p_1, p_2, \cdots, p_n)$ between those two points so that $p_1 = p$ and $p_n = q$ and for each i, $\|I(p_i) - I(p_{i+1})\| < \lambda$*

$\|.\|$ is a L_2-norm and increasing values of λ create a hierarchy of partitions.

The watershed is a region growing algorithm which defines a pathwise connection. The watershed lines associate a catchment basin to each minimum of a function. Typically, the function to flood is a gradient function which catches the transitions between the regions. Region seeds of the watershed are therefore the gradient minima. The waterfall algorithm [5] enables to construct a non parametric hierarchy of watersheds (a hierarchy of partitions) which performs region merging between adjacent catchment basins. Flat zones hierarchies usually are too fine and waterfall ones too coarse, thus we propose a new connective criterion which is an intermediate one between pathwise connections (watersheds) and threshold connections (quasi flat zones). This new criterion is referred to as homogeneous connections.

Definition 4 *Two points p and q belong to a same homogeneous zone of an image I if $\|I(p) - I(q)\| \leq k \times \lambda(Seed(p))$ with $Seed(p)$ the initial seed of the region of p and $\lambda(p) = \frac{1}{n_v} \sum_{p_v \in V(p)} \|I(p) - I(p_v)\|$*

$V(p)$ denotes the neighbors of p and n_v the cardinal of this set, $\|.\|$ is a L_2-norm and k is a real number which sets the fineness of the partition. $\lambda(p)$ being close to a gradient computation, pixels in homogeneous regions (the color variation among the considered neighborhood is small) will be considered first as candidate region seeds. Each pixel is a candidate region seed which grows by aggregating adjacent pixels according to the previous rule. This implies that a pixel q is aggregated to a region R if the distance between a pixel p of R, neighbor of q, is k times lower than the initial homogeneity of the seed pixel of R. k is the accepted homogeneity jump and states if two pixels belong to the same region. Homogeneous zones therefore produce partitions, the fineness of which decreases while k increases (the homogeneity constraint is slackened). Obviously, a hierarchy of partitions obtained for increasing values of k is not nested since it lowers the number of initial seeds while slackening the homogeneity constraint. This is the same problem as producing nested partitions with the watershed: one has to consider the output of the i^{th} level as an input for the $(i+1)^{th}$ level. Therefore it is possible to produce hierarchical partitions using homogeneous connections by applying the same principle on the partition obtained at the previous level (an efficient implementation uses graphs).

Homogeneous connections produce hierarchical partitions which are finer than the waterfall and coarser than the quasi flat zones. However they are better suited in the case of automatic segmentation since they do not need a definition of seeds on a given gradient and they locally adapt their behavior to the image content since the threshold which determines if a pixel belongs to a region depends on the local homogeneity at this pixel. Figure 1 presents several levels of hierarchical partitions obtained by quasi flat zones, homogeneous connections and waterfall. The saliency map (right column) illustrates the importance of each pixel among the levels, the saliency of a pixel being defined as the highest level for which it occurs at the boundary between two regions (in the saliency map images, the gray level corresponds to the hierarchy level, therefore the highest the brightest).

Partitions obtained from any connective criterio, provide, in the low levels, over-segmented partitions called fine partitions. These fine partitions are interesting for image simplification because they can be used as markers for connective filters such as levellings [6]. However they have another strong interest. Hierarchies of partitions, such as irregular pyramids [7], usually use very simple fine partitions as the lowest level of the pyramid, namely each pixel is a region. It is totally useless to proceed in such a way since it makes the hierarchical structure bigger by including evident mergings at the lowest levels. Moreover the pixel grid is not a natural representation of visual scenes. It is much more natural, and presumably more efficient, to work with perceptually meaningful entities obtained from a low-level grouping process. To build a hierarchy of partitions, one thus might use a fine segmentation, obtained from any connective criterion, as the base of the pyramid. In this paper we are interested in building such hierarchies.

Figure 1: Hierarchies of partitions by quasi flat zones (top), homogeneous connections (middle) and waterfall (bottom). Depicted partitions correspond the levels 1, 5, 10 and the last column presents the saliency maps of the produced hierarchies over ten levels.

3. MULTI-SCALE IMAGE SEGMENTATION

In the previous section we showed several connective criteria which can be used to produce hierarchical partitions. The produced partitions can be considered, for the finest levels, as a low-level grouping process subject to further analysis to generate an accurate multi-scale segmentation in a hierarchy of partitions. Besides connective criterion, scale-space and region merging can also be used to generate hierarchical partitions. Scale space approaches for hierarchical segmentation use a scale generation and a linking mechanism. The scale generation is usually performed by a diffusion and the linking scheme aims at tracking the regions through the scale-space [8]. The linking scheme is essential since without it the inclusion relationship between successive levels of the scale-space segmentation stack is not preserved. This is not the case with region merging approaches but the difficulty is reported on a region merging predicate which states if two regions have to merge or not. In this section we propose to combine these two approaches into a single one to produce efficient hierarchical partitions.

3.1 Graph simplification

A fine partition is over-segmented but is very close to the content of the original image. Constructing the so-called mosaic image (each pixel has the average color of its surrounding region) is equivalent to perform an image simplification the result of which is piecewise constant. As stated before there is no interest in working directly on the pixel grid and it is more interesting to work on a fine partition. Once a partition P_k is considered, an equivalent representation is its Region Adjacency Graph (RAG) G_k, a hierarchy of partitions then becomes a stack of nested graphs. A RAG is a set of nodes representing connected components (the regions) of the image and a set of links or edges E connecting two neighboring nodes. This RAG denoted by $G = (V, E)$ is constructed to describe a partition of the image by the topology and the inter-region relations of the image. It is defined by an undirected graph where $V = \{1, 2, ..., K\}$ is the set of nodes (Vertices) and $E \subset V \times V$ is the set of edges (links between adjacent regions). $K = |G|$ is the number of region nodes. To each node is associated a model corresponding to the region the node represents, we consider a very simple model: the average color of the region. Visual objects in an image being significant only at a given scale level, this is also the case for the nodes of the RAG. We can formalize this aspect by scale-space approaches, namely consider the image at different scale levels. Contrary to classical scale-space approaches which directly operate on the pixel grid image [8], we propose to generate successive scales directly on a RAG obtained from a fine partition, and this comes to simplify the models associated to the nodes. We propose to perform this simplification by a non linear graph model simplification which simplifies the models attached to nodes and therefore generates a set of images corresponding to successive simplification scale levels. To simplify the RAG, an iterative process is used. Given the initial RAG ($t = 0$), a new model is computed for each node at each iteration ($t > 0$) according to the neighbors of each node. This new model is defined by the following expression:

$$V_i^{t+1} = \frac{\sum_j \left(\alpha_{ij}^t \cdot V_j^t\right)}{\sum_j \alpha_{ij}^t} \qquad (1)$$

with $V_i^0 = V_i$, $\alpha_{ij}^t = g\left(d(V_i^t, V_j^t)\right)$ where the V_j^t node is a neighbor of the V_i^t node in the RAG at the iteration t. V_i^t denotes the mean color value associated to a node of the graph. d is the classical euclidean distance between two color distributions. g is a weighting decreasing positive function which is defined in this paper by $g(s) = e^{-\left(\frac{s}{k}\right)^2}$. This approach is similar to the one used by PERONA and MALIK in the choice of their function for anisotropic diffusion [9] often used to generate scale-space stack of images [8]. The graph simplification defined here produces a new simplified RAG where the model associated to each node (a region in the image partition) is obtained according to its neighbors. A first simple output of this method used alone is that it produces a simplified image (mosaic image)when associating to each pixel of the original image the color of its surrounding region (a node of the graph). Figure 2 presents an example of image simplification. The initial fine partition was built from strict flat zones with $\lambda = 0$ (the mosaic image is exactly the initial image) and the mosaic images of the simplified RAGs are obtained after 0, 5, 15, 50 and 200 iterations of the graph simplification. Bottom row presents results on the same image corrupted by gaussian noise. The simplified images could have been obtained by a similar simplification on the initial image, however performing the simplification on a RAG enables a considerably faster execution time. This is interesting for several reasons. Firstly this method is a faster alternative to classical image simplification since it operates on a set which is much less important than the whole pixel grid. Secondly classical scale-space generation by image simplification implies a loss of resolution and a displacement of edges across the scales which has to be solved by the linking scheme, this is not the case with our approach. Thirdly the obtained RAG can be used to obtain easily an improved segmentation as compared to the original fine one. Therefore we propose to couple graph simplification with graph decimation and to perform merging of nodes after each RAG simplification step. Since the model attached to each RAG is simplified at each iteration, similar regions tend to similar models and can be merged. This will decrease once again the computation time since the simplification will operate on a restricted RAG after the merging of similar nodes. This is the core of the next section.

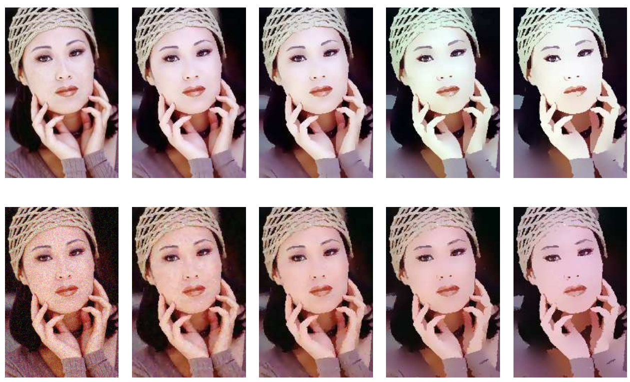

Figure 2: Set of simplified images in a scale-space approach on RAG after 0, 5, 15, 50 and 200 iterations (from left to right) on the original image (top row) and a gaussian noise corrupted one (bottom row).

3.2 Graph decimation

The graph simplification method that we proposed in the previous section is right to obtain simplified versions of the models associated to nodes of the RAG of a fine partition. However the simplification being a kind of scale-space on graph, it does not generate a hierarchy of partitions but a hierarchy of images simplified at different scale levels. We can nevertheless take advantage of the simplification to simplify the structure of the graph too. In fact, image simplification tends to bring similar models closer and similar regions can merge then. The idea of merging regions in a partition is quite old [10] and is the basis of a lot of hierarchical methods such as irregular pyramids [7]. For a complete merging strategy based on a RAG, several notions have to be defined [11]:

- The region model M_R: a model defines how to represent a region and also the union of two of them.
- The merging order $O(R_i,R_j)$: it associates to each edge of the RAG a similarity measure (a weight) between adjacent nodes. This order is a function defined for each couple of neighbor regions and its values belong to a totally ordered set Λ which provides the set of scales.
- The merging predicate $C(R_i,R_j)$: this criterion defines if two regions have to merge or not.

Creating a hierarchy of partitions by a region merging algorithm simply consists in pairwise merging of regions and in updating the RAG structure [12]. For each threshold $\lambda \in \Lambda$, one can define a contraction kernel [2] on the graph which merges regions the edge weight of which is lower than a given λ threshold. This provides a partition P_λ for each scale λ. The construction of P_λ is equivalent to finding the maximal connected components on the graph the similarity of which is under the scale level λ and $H = \{P_\lambda\}_{\lambda \in \Lambda}$ is a hierarchy of nested partitions since each region of $P_{\lambda+1}$ is a disjoint union of regions of P_λ. We propose to combine this type of hierarchical segmentation which proceeds to a graph decimation with the previously proposed graph simplification. The principle is iterative and consists in alternating simplification and decimation of the RAG. At each iteration, models attached to each node are simplified and similar regions are merged according to a merging criterion [13]. As for simplification, the region model is very simple and is the average color of each region: $M_{R_i} = V_i^t$. The union of regions having to be computed fast, it is defined directly from the two models merging: $M_{R_i} \cup M_{R_j} = V_i^t \cup V_j^t = V_i^t + V_j^t$. The merging order is directly based upon the similarity between regions such as it was defined for the graph simplification: $O(R_i,R_j) = O(V_i^t,V_j^t) = d(V_i^t,V_j^t)$. To perform the merging of regions fulfilling the merging criterion, edges are ranked into a hierarchical priority queue according to the edge weights. At each merging, the edge of minimum cost is removed from the hierarchical queue, the region model of the merging and the weights of all the edges of the adjacent regions computed, some edges being suppressed. The whole algorithm is depicted in algorithm 1 where λ is the level of the hierarchy, one goes from one scale level λ to the next one $(\lambda+1)$ only if regions have merged. $|.|$ denotes the cardinality of a set.

```
λ : integer; λ_end : integer;
λ ← 0; Define λ_end
G_λ = (V_λ, E_λ) for an initial partition P_λ.
While (λ ≤ λ_end) do
    For the nodes V_i ∈ V_λ do
        Simplify the node model V_i.
    end For
    For the edges E_l = (V_i, V_j) ∈ V_λ × V_λ do
        If (C(V_i, V_j)) then
            Add E_l to the contraction kernel CK_{λ,λ+1}
        end If
    end For
    Contract the graph G_λ with the contraction kernel
    CK_{λ,λ+1} : G_{λ+1} = Contraction[G_λ, CK_{λ,λ+1}]
    If (|CK_{λ,λ+1}| > 0) then
        λ ← λ + 1
    end If
done
```

Algorithm 1: Algorithm for multi-scale image segmentation in a hierarchy of partitions by graph simplification and decimation.

However one thing remains to be defined: the merging predicate $C(V_i,V_j)$. In this paper we consider several possible choices: a fixed threshold, an evolving threshold and an adaptive threshold. For fixed threshold the merging predicate is $O(V_i,V_j) < T$. For an evolving threshold, the merging predicate is the same but after each iteration the threshold is increased $T = T + \Delta T$, $\Delta T = 0.5$ in this paper. For an adaptive threshold we use the formulation of Nock (see in [14]). Figure 3 presents multi-scale segmentations obtained by the proposed approach with the different merging predicates for the levels 1, 4, 9, 15 and 20: fixed threshold equal to 1 (rows 1 and 2), evolving threshold (rows 3 and 4) and adaptive threshold (rows 5 and 6). For each hierarchy, the simplified image and the partition are provided. The initial partition was obtained by homogeneous connection ($k = 0.5$, row 7, first image). Saliency maps are given in row 7 for the different thresholds. With a fixed threshold the number of scales is high and the image simplification produced is very close to the original image even if the number of regions is much less important than the initial number of pixels. For the other predicates, less levels and coarser segmentations are obtained faster. This is related to the combination of simplification and decimation when the decimation criterion changes across scales. The produced hierarchies are good and extract the primary visual components from the image. Contrary to several multi-scale methods which first simplify the image and then segment it [15], our approach enables to combine these two approaches in a single faster algorithm.

4. CONCLUSION

In this paper, we proposed an algorithm for multi-scale image segmentation in a hierarchy of partitions. This algorithm couples graph simplification and decimation based on non linear smoothing and region merging. Moreover the base level of the hierarchy is obtained by generating a fine partition by a new connective criterion. The proposed approach blends together into a single algorithm scale-space and region merging for multi-scale image segmentation without the need of a linking scheme to follow regions in the scale-space stack. Obtained hierarchies are very good and the only parameter of the method is the merging predicate which can be fixed, evolving or adaptive. The algorithm can be used for image simplification as well as for image segmentation, both being multi-scale.

REFERENCES

[1] J-M. Jolion and A. Montanvert. The adaptative pyramid, a framework for 2d image analysis. *Computer Vision Graphics and Image Processing*, 55(3):339–348, 1992.

[2] W. Kropatsch. Building irregular pyramids by dual graph contraction. In *IEE Proceedings of Vision, Image and Signal Processing*, volume 142, pages 366–374, 1995.

[3] J. Serra. Connectivity for sets and functions. *Fundamenta Informaticae*, 41:147–186, 2002.

[4] F. Meyer. From connected operators to levelings. In *Proceedings of ISMM'98*, pages 191–199, 1998.

[5] S. Beucher. Watershed, hierarchical segmentation and waterfall algorithm. In *Proceedings of ISMM'94*, pages 69–76, 1994.

[6] J. Angulo and J. Serra. Morphological coding of color images by vector connected filters. In *Proceedings of ISSPA*, volume 1, pages 69–72, 2003.

[7] P. Meer. Stochastic image pyramids. *Computer Vision, Graphics and Image Processing*, 45(3):269–294, 1990.

[8] I. Vanhamel, I. Pratikakis, and H. Sahli. Multiscale gradient watersheds of color images. *IEEE transactions on Image Processing*, 12(6):617–626, 2003.

[9] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[10] S.L. Horowitz and T. Pavlidis. Picture segmentation by a tree traversal algorithm. *J. Assoc. Compt. Math.*, 2(23):368–388, 1976.

[11] 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.

[12] S. Makrogiannis, G. Economou, and S. Fotopoulos. A graph theory approach for automatic segmentation of color images. In *International Workshop on Very Low Bit-rate Video*, pages 162–166, 2001.

[13] P. Felzenswalb and D.P. Huttenlocher. Efficient graph-based image segmentation. *IJCV*, 59(2):167–181, 2004.

[14] R. Nock and F. Nielsen. Statistical region merging. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 26(11):1452–1458, 2004.

[15] B. Sumengen and B. S. Manjunath. Multi-scale edge detection and image segmentation. In *European Signal Processing Conference (EUSIPCO)*, Sep 2005.

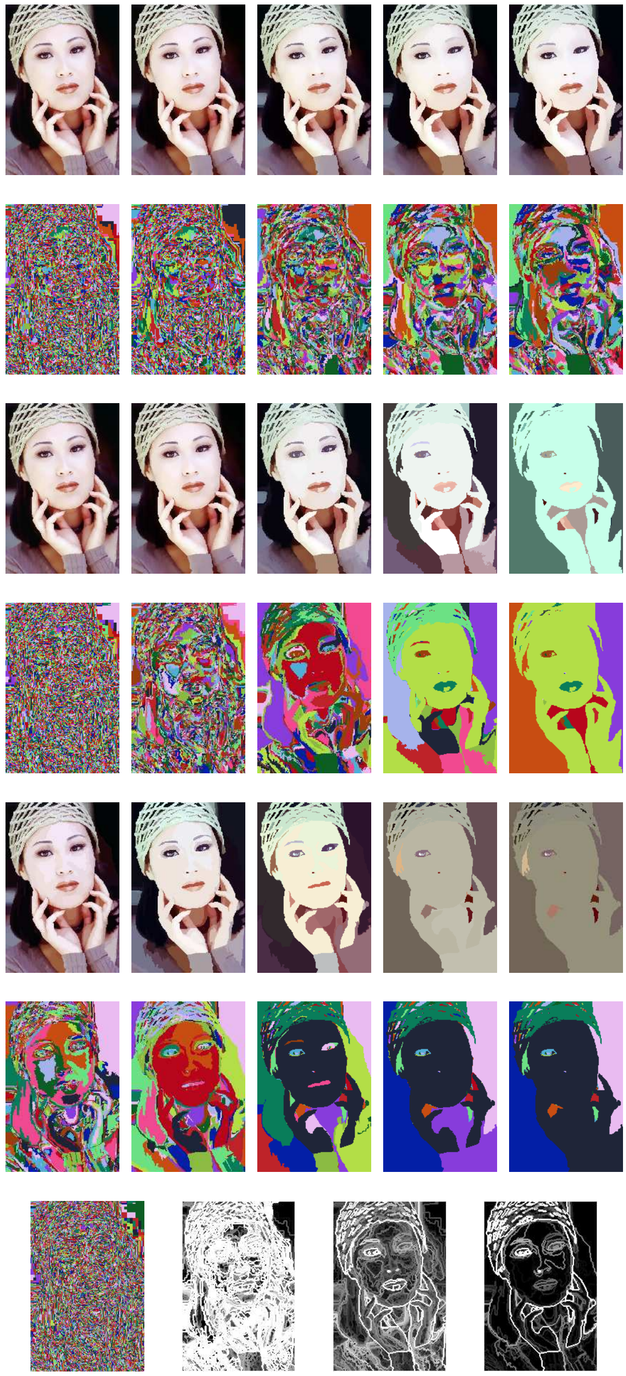

Figure 3: Multi-scale segmentations in a hierarchy of partitions for scales 1, 4, 9, 15 and 20 with different merging predicates (see text).