An unsupervised color image segmentation based on morphological 2D clustering and fusion

O. Lezoray
LUSAC EA 2607, Vision and Image Analysis Group
IUT SRC, 120 Rue de l'exode, 50000 Saint-Lô, FRANCE
Olivier.Lezoray@info.unicaen.fr
Tel: +33(0)233775517
Fax: +33(0)233771167

Abstract

A segmentation method for color images is presented in this work. A morphological unsupervised 2D multiband histogram clustering provides an initial coarse segmntation of the image. Region information is then used and a novel technique is introduced to simplify the Region Adjacency Graph by merging candidate regions until the stabilization of a segmentation criterion. Merged regions are refined by a color watershed. The whole method requires the definition of only one single parameter which acts as a kind of multiscale parameter.

1. Introduction

Color image segmentation refers to the partitionning of a multi-channel image into meaningful objects. With the growing of digital image databases, efficient segmentation methods are needed for extracting and coding image regions. Various approaches can be found in the litterature and can be roughly classified into several categories: clustering methods [1], edge-based methods [2], region growing methods [3], morphological watershed based region growing methods [4]. Many of the existing segmentation techniques, such as supervised clustering use a lot of parameters which are difficult to tune to obtain a segmentation where the image has been partitionned into homogeneous colored regions. We propose an unsupervised approach which minimizes the number of parameters to one multi-scale parameter. To this aim, we combine different types of methods to obtain a segmentation of a color image. The basic idea of is to divide the segmentation process into three stages [5] : color clustering, region merging and watershed segmentation. In the first stage, 2D histograms are used to obtain a rapid and coarse clustering of the color image. This clustering is fast, simple and unsupervised (the number of classes is automatically estimated with morphological operators). This segmentation is however over-segmented and the second stage proceeds to a region merging of adjacent regions until the stabilization of a cost associated with the partitionning of the color image. In the third stage, a segmentation refinement is based on a morphological filtering and a color watershed. The three-stages segmentation enables to consider in the first stage the spatial distribution of the colors to cluster the image and in the next two stages the color similarity between regions and pixels or regions. The whole segmentation scheme is performed in the $L^*a^*b^*$ color space for its uniform perceptual properties.

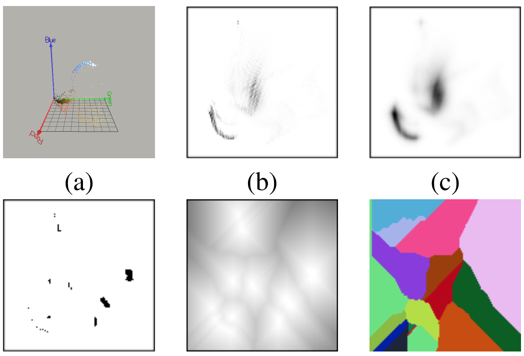
2. 2D multiband histogram

2.1. Clustering of color histograms

To perform a clustering of a color image, several strategies can be employed. The strategies generally differ on the dimension of the data used to cluster the image. In the 1D case, the histograms of each band are considered separately and the method is reduced to finding thresholds dissecting the band histograms [6, 7]. This method relies on the fact that the objects in the image give rise to explicit peaks in the histograms. The resulting segmentation maps can be combined with different methods, such as majority vote, Demspter-Shafer or Bayesian theory. This method is known as multithresholding. In the 3D case, the vectorial aspect of color is taken into account and the clustering is considered as the classification of multispectral data. The spatial repartition of the colors in the associated color space is used to cluster the image assuming that the colors of the objects are grouped around dominant colors in the 3D histogram [8, 9, 10, 11] (the center of the classes). Both 1D and 3D methods suffer from the determination of the number of classes, which is usually assumed to be known. The 1D method also suffers from the fact that a color cluster is not always present in each band and the combination of the different segmentations cannot catch this spatial property of the colors. The 3D method is handicapped with data sparseness on the one hand and with the complexity of the search algorithm on the other hand. An interesting alternative to these two methods lies on the use of partial histograms [12] (2D histograms) which use two color bands together namely L^*a^*, L^*b^* and a^*b^* in the $L^*a^*b^*$ color space. This can bring several advantages. The paucity of the data encountered in the 3D case is partially overcame and the search complexity is drastically reduced. Moreover it partially uses the spatial repartition of color and offers an intermediate method to the 1D and 3D ones. Another advantage to be considered is the fact that a 2D histogram is nothing more than a grey-level image, therefore classical and fast grey-level image processing algorithms can be used to cluster the 2D histogram.

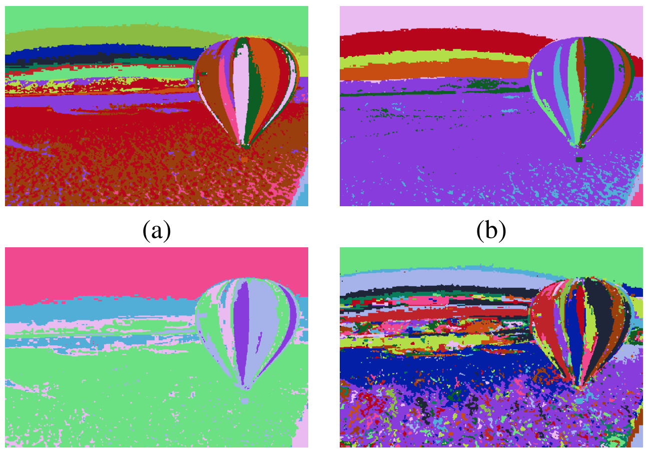
2.2. Unsupervised Morphological clustering

In this section we propose a new unsupervised morphological method for clustering 2D histograms obtained from color images. A 2D histogram is the projection of a 3D histogram onto band-pair planes (see figure 1(a)-(b)). To cluster this 2D histogram, we assume that the different objects of the image are present in the histogram around the dominant peaks. The main peaks of the 2D histogram are considered as the cluster centroids. Since the 2D histograms are generally noisy due to the sparseness of the colors in the images, the latter is smoothed with a symetrical exponential filter (with a given parameter $\beta \in [0, 1]$). The result is reconstructed in the original image to obtain a simplified version of the histogram (see figure 1(c)). The main peaks corresponding to maxima in the 2D histogram, an erosion applied on the latter reduces each bin to its main colors. The dominant colors therefore correspond to the ultimate erode set of the 2D histogram (see figure 1(d)). This method directly extracts the color cluster centroids whithout any assumption on the number of classes. Using these centroids as markers, a watershed is performed on the the distance function to the markers and this provides the clustering of the histogram (see figure 1(e)-(f)). From the clustered 2D histogram a segmentation map is obtained since each region in the histogram image corresponds to a set of colors in the original image. The above clustering method is applied to the three 2D histograms of a color image and three different segmentation maps are obtained (see figure 2(a)-(c)). The merging of the three different segmentation maps is deduced from their intersection. Finally a label is then given to each region of the segmentation map and a majority filter is applied to the resulting image to suppress isolated pixels. One can state that the final regions extracted (see figure 2(d)) correspond to homogeneous regions in the original color image (see figure 5(a)).



(a) (b) (c)
(d) (e) (f)

Figure 1: *(a) The 3D histogram in the $L^*a^*b^*$ color space (b) The L^*a^* 2D histogram (inverted). (c) The simplified histogram (inverted) (d) The dominant colors (e) The distance function to the dominant colors (f) The watershed clustering.*



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

Figure 2: *(a)-(c) the segmentation maps of the L^*a^*, L^*b^*, a^*b^* 2D histograms (d) The intersection of the three segmentation maps after a majority vote filtering.*

3. Region Merging

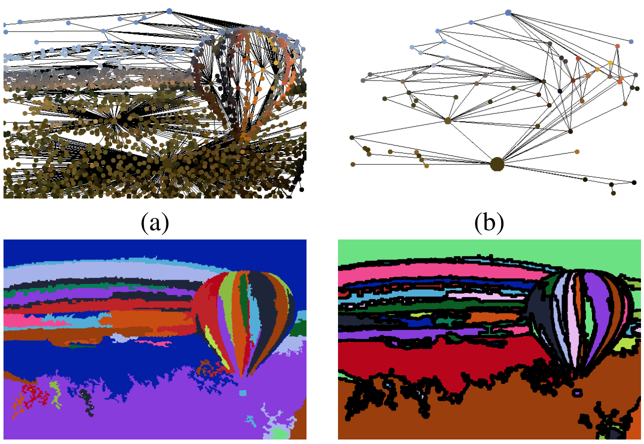
Given the initial image clustering which defines an image segmentation (often oversegmented), a merging strategy is needed to join the most coherent adjacent regions together. We propose to use a Region Adjacency Graph to simplify the initial image segmentation by merging adjacent regions.

3.1. RAG Construction

A RAG is a set of nodes representing connected components (the regions) of the image and a set of links connecting two neighboring nodes [13]. 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 and $E \subset V \times V$ is the set of edges (links between adjacent regions). $K = \Theta_0(G)$ is the number of region nodes. Each edge is weighted by a value indicating the similarity between two adjacent regions. The similarity between two regions (the weight of the edge) R_1 and R_2 is defined by the following expression [13]:

$$E(R_1, R_2) = N_1 \|M_{R_1} - M_{R_1 \cup R_2}\|_2 + N_2 \|M_{R_2} - M_{R_1 \cup R_2}\|_2$$

where N_1 and N_2 are the number of pixels of the regions R_1 and R_2. $M(R)$ is the region model (the mean color) and $\|.\|_2$ is the L_2 norm. This similarity defines the order in which the merging of the regions has to be performed. The most similar pair of adjacent regions corresponds to the edge with the minimum cost. A merging algorithm on the RAG uses this property to merge region nodes corresponding to edges of minimum cost. A merging algorithm on a RAG is therefore a technique that removes links and merges the corresponding nodes. The algorithm is simple: the pair of most similar regions are merged until a termination criterion is reached.



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

Figure 3: (a) Initial RAG (b) Simplified RAG (c) Regions after merging (d) Morphological filtering.

3.2. Merging Termination Criterion

The termination criterion is a crucial point of the algorithm and defines when the merging ends. For an automatic segmentation, the termination criterion should be based on image statistics. To assess this fact, we have defined a new criterion (based on the one developed by Liu [14]) using the RAG structure, the merging order and the image model. The merging of similar regions is performed until a stabilization of the segmentation evaluation criterion (successive identical values). The segmentation evaluation F is defined by the following expression:

$$F(G, I) = \frac{\sqrt{\Theta_t(G)}}{1000 \times h \times w} \times \left(\sum_{i=1}^{\Theta_t(G)} \frac{e_i^2}{1 + log(N_i)} + \sum_{j=1}^{\Theta_t(G)} \sum_{k>j}^{\Theta_t(G)} E_{j,k} \right)$$

where G denotes the graph and $\Theta_t(G)$ denotes the number of nodes in the graph at a given iteration t : this number decreases along the iterations. I is the initial color image, h, w are respectively the height and the width of I. e_i^2 is the euclidean distance between the $L^*a^*b^*$ color vectors of the pixels of the ith region in the original image I and the mean color vector attributed to the ith region in the segmented image. If the segmentation is enough representative, each region has an homogeneous mean color as compared to the original color vectors, this is equivalent to an intra class distance. Moreover the merging of two adjacent regions makes smaller the sum of the edge weights, this is somehow equivalent to an inter class distance. Once the segmentation evaluation function appears to be stabilized, a trade-off between the inter and intra class distances has been reached. An example of merging is given by the figure 3. To accelerate the processing, the computing of e_i is performed in the following way. If two regions R_1 and R_2 merge then

$$e^2_{R_1 \cup R_2} = \frac{N_{R_1} e^2_{R_1} + N_{R_2} e^2_{R_2}}{N_{R_1 \cup R_2}}$$

This enables a faster implementation of the computing of $F(G, I)$ which is rapidly updated each time two regions merge. A plateau is considered to be reached when a certain number of successive identical values of F is obtained. This number is dependent on the number of regions in the original segmentation map and is defined by the quantity $\frac{\sqrt{\Theta_0(G)}}{2}$ where $\Theta_0(G)$ gives the initial number of regions before the merging process begins. The figure 4 illustrates the variation of the F segmentation evaluation criterion. It decreases along the iterations until it reaches a plateau ending the merging process. An iteration corresponds to the analysis of an edge link between two regions.

4. Color Watershed refinement

Unfortunately the segmentation obtained after the RAG simplification can produce irregular boundaries of the objects. To reduce this effect, a morphological filtering is

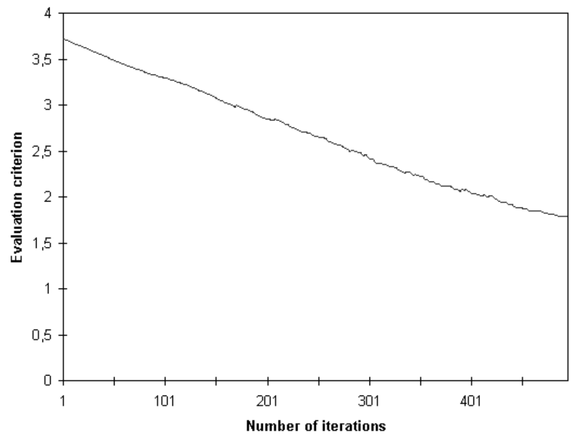


Figure 4: *Variation of F along the iterations*

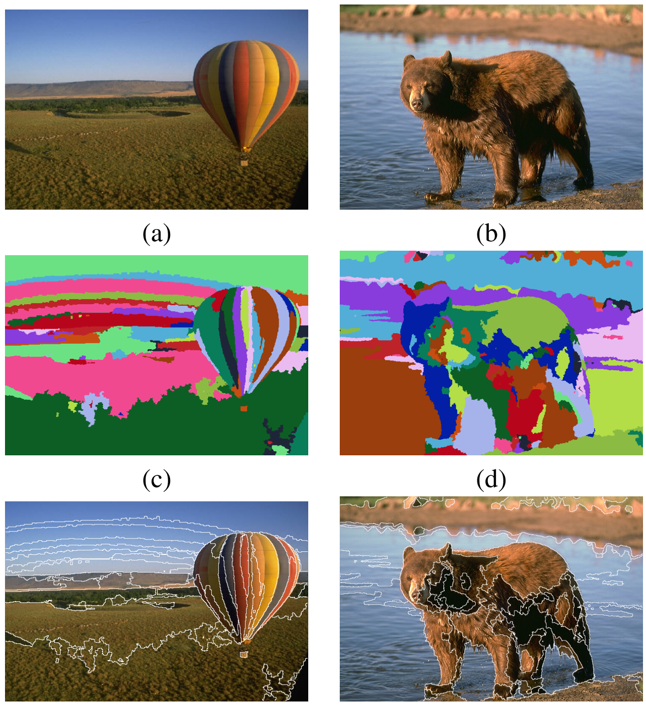
applied. Two sucessive label erosions are performed: if a pixel has a label j and any of its neighbors has a label $k \neq j$ or is unassigned, then the pixel is set to unassigned (see Figure 4(d)). Markers are set to all the remaining regions after the morphological filtering and a color watershed algorithm [5] is then used to fill the unassigned areas.

5. Results and Conclusion

The algorithm proposed in this paper has been tested over a large number of color images from the corel Image CD Database. Figure 5(a)-(f) shows the segmentation results for a small sample set. All the experiments were led with an experimentally fixed value of the single parameter of the method ($\beta = 0.4$). To assess the influence of the value of the single parameter β, we have to compare the obtained segmentations for each value of β. To see if a segmentation is close to the original image, an error metric is needed. The error between the original image and the quantized image (obtained by associating the main color of each region to each segmented pixel) is generally used. The Mean Square Error (MSE) is therefore considered to evaluate a segmentation. The MSE is defined by

$$MSE = \frac{1}{h \times w} \sum_{i=1}^{h} \sum_{j=1}^{w} e_i^2$$

A lower value for MSE means lesser error, therefore the lower the MSE, the better the segmentation. The figure 6 presents the MSE and the number of regions for each value of β. The number of obtained regions obviously decreases with β while the MSE increases. One thing to point out is the fact that β is not a strict multi-scale parameter since the number of regions obtained with $\beta = 0.9$ is lower than with $\beta = 0.8$. This is mainly due to the region merging



(a) (b) (c) (d) (e) (f)

Figure 5: *Two segmentations after the color watershed ($\beta = 0.4$ in the $L^*a^*b^*$ color space): (a)-(b) initial images (c-d) final regions (e-f) superimposed boundaries.*

step: the more the number of regions in the original segmentation to simplify, the more the size of the plateau of F to be reached. The final segmentation therefore does not fullfill the requirements of a real multi-scale parameter where the number of regions has a monotonous decrease with β. One can however state that a "good" value for β is between 0.3 and 0.5. This parameter might be however determined for each image to be processed to enable the adapdation of the method. The filtering of the $2D$ histogram will therefore be adapted to the spatial repartition of the colors in the histogram. One can even consider different values of β for each $2D$ histogram.

References

[1] E. Littman and E. Ritter, "Colour image segmentation : a comparison of neural and statistical methods," *IEEE Transaction on Neural Networks*, vol. 8, no. 1, pp. 175–185, 1997.

[2] D. Zugaj and V. Lattuati, "A new approach of color images segmentation based on fusing region and edge segmentation outputs," *Pattern Recognition*, vol. 31, no. 2, pp. 105–113, 1998.

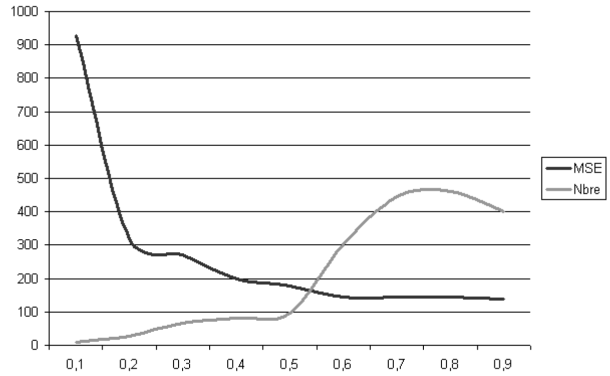


Figure 6: The MSE and the number of regions according to the value of β for the segmentation of the image 5(a)

[3] A. Trémeau and N. Borel, "A region growing and merging algorithm to color segmentation," *Pattern Recognition*, vol. 30, no. 7, pp. 1191–1203, 1998.

[4] L. Shafarenko, M. Petrou, and J. Kittler, "Automatic watershed segmentation of textured color images," *IEEE transactions on Image Processing*, vol. 6, no. 11, pp. 1530–1543, 1997.

[5] O. Lezoray and H. Cardot, "Hybrid color image segmentation using 2d histogram clustering and region merging," in *ICISP'03*, 2003, vol. 1, pp. 22–29.

[6] M. Celenk, "A color clustering technique for image segmentation," *Computer Vision Graphics and Image Processing*, vol. 52, pp. 145–170, 1990.

[7] O. Lezoray and H. Cardot, "Histogram and watershed based segmentation of color images," in *Proceedings of CGIV'2002*, 2002, pp. 358–362.

[8] P. Soille, "Morphological partitioning of multispectral images," *Journal of Electronic Imaging*, vol. 18, no. 4, pp. 252–265, 1996.

[9] S.H. Park, I.D. Yun, and S.U. Lee, "color image segmentation based on 3-d clustering: morphological approach," *Pattern Recognition*, vol. 31, no. 8, pp. 1061–1076, 1998.

[10] J.G. Postaire, R.D. Zhang, and C. Lecocq-Botte, "Cluster analysis by binary morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 2, pp. 170–180, 1993.

[11] T. Géraud, P.Y. Strub, and J. Darbon, "Color image segmentation based on automatic morphological clustering," in *IEEE International Conference on Image Processing*, 2001, vol. 3, pp. 70–73.

[12] F. Kurugollu, B. Sankur, and A. Harmanci, "Color image segmentation using histogram multithresholding and fusion," *Image and Vison Computing*, vol. 19, no. 13, pp. 915–928, 2001.

[13] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation and information retrieval," *IEEE transactions on Image Processing*, vol. 9, no. 4, pp. 561–576, 2000.

[14] J. Liu and Y-H. Yang, "Multiresolution color image segmentation," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 689–700, 1994.

Biography

O. LEZORAY has received has the M.Sc. degree in Computer Science in 1995 and the Ph.D. degree in computer Science from the University of Caen in 2000. From September 2000 to august 2001 he worked as a one year assistant lecturer in the Computer Science department of the University of Caen. Since September 2001, he joined the Communication Networks and Services department of technology of the University of Caen as a senior lecturer. His research focus on image segmentation techniques for color images and data classification methods based on the cooperation of machine learning methods.