Unsupervised Hierarchical Spatio-Colorimetric Classification For Color Image Segmentation

Cindy Torres, Alain Clément
Université d'Angers, Laboratoire d'Ingénierie des Systèmes Automatisés (EA 4014), Institut Universitaire de Technologie, 4 Bd Lavoisier, BP 42018 - 49016 Angers Cedex. France.
cindytorres@free.fr, alain.clement@univ-angers.fr

Abstract

A new unsupervised vectorial segmentation method developed for color images is presented. Both spatial and color information have been used during the classification process of the pixels. To overcome the problem of memory space associated with multidimensional histograms analysis and avoid a color requantization, this method is based on a multidimensional compact histogram and an original compact spatial neighborhood probability matrix. The multidimensional compact histogram allows a drastic reduction of memory space necessary for coding color histograms without any data loss. Leaning upon the compact histogram, a spatial neighborhood probability matrix has been computed. It contains all non-negative probabilities of spatial connectivity between pixel colors and allows a spatial distance inside a variable size neighborhood to be defined. In an unsupervised histogram analysis classification process, two phases are classically distinguished: (i) a learning process during which histogram modes are identified and (ii) a second step called the decision step in which a full partition of the colorimetric space is carried out according the previously defined classes. In a standard colorimetric approach, in the second step, a colorimetric distance like Euclidean or Mahalanobis is used. We introduced here a spatio-colorimetric distance taking into account the information of pixel neighborhood colors. This distance is defined as a weighed mixture between a colorimetric distance and the spatial distance calculated from the spatial neighborhood probability matrix. The vectorial classification method is based on previously presented principles, achieving a hierarchical analysis of the color histogram using a 3D-connected components labelling.

Introduction

Segmentation is an important step in the image processing chain for identifying and partitioning the different regions of interest in an image. Classically, the algorithms for segmenting images can be divided into two families: the ones using the image plane spatially and the others using the color distribution of the pixel in the selected color space.

In color images, a pixel is considered as a three-dimensional (3D) vector whose components depend on the color space used. When color distribution is chosen, it is supposed that colors of homogenous regions give rise to clusters in the color space, each of them corresponding to a class of pixels. The different classes are obtained by a cluster analysis or by means of a color mode detection method generally based on color histograms. This procedure assigns each pixel of the image to a class depending on its color. By connecting pixels from the same class, regions are constituted.

The entire color information is frequently not used because of the computation time required. That is why, most color classification methods are based on mono or bi-dimensional histograms regardless of the correlation between color planes is lost. Trémeau and Laget [1] have demonstrated however, by the Shannon theory, that using a multidimensional analysis is more discriminatory than a mono-dimensional analysis. To overcome the problem of memory space Xuan [2] proposes to requantify the color scale by reducing the number of bits for coding each color component. This technique is efficient, but it performs an a priori color classification.

Among segmentation methods using color space classification, the ones relying on histograms analysis have the advantage of being unsupervised but have also the drawback of not taking into account the spatial information. For the last few years, a third family of segmentation methods has appeared: spatio-colorimetric classification methods.

For some images, the loss of the spatial information conducts to a false segmentation. That problem was highlighted by Trémeau [3] and Macaire et al. [4]. With natural color images, the classification generally leads to an over-segmented image with small regions scattered through the image. This may be explained by the lack of correspondence between some peaks in the color space and significant regions in the image, or by the merging of too small peaks with higher ones colorimetrically close but corresponding to inhomogeneous spatial regions. To cope with these problems, original approaches taking into account both spatial connectivity and color information have been proposed by Macaire et al. [4] Foucher [5] Trémeau [3] Busin et al. [6] Noordam and Broek [7] Comaniciu and Meer [8]. They have developed original supervised or not algorithms or fixed important axioms.

These approaches are facing the double difficulty of treating a huge quantity of information and dealing with a high algorithmic complexity.

In this paper a new contribution to unsupervised spatio-colorimetric classification is presented. For several years, our laboratory has been involved in developing classification algorithm based on multidimensional histograms [9] [10]. Thanks to the compact histogram [11] and an original compact spatial neighborhood probability matrix, a new unsupervised vectorial segmentation method taking into account the full 3D histogram and the spatial organization of pixels has been developed. This method is based on a hierarchical analysis of the histogram. In a standard colorimetric approach, colorimetric t-uples are attributed to classes minimizing a colorimetric distance like Euclidean or Mahalanobis. We insert here a spatio-colorimetric distance taking into account the information of pixels neighborhood colors. This distance is defined as a weighed mixture between a colorimetric distance and the spatial distance calculated from the spatial neighborhood probability matrix. The vectorial classification method is based on the spatio-colorimetric distance and achieves a hierarchical analysis of the color histogram using a 3D-connected components labeling.

In a first part, the principle of the compact histogram is explained and the spatial neighborhood probability matrix is detailed.

Secondly, the hierarchical unsupervised classification method is presented and the spatio-colorimetric distance is defined.

In a third part, the classification method has been applied to synthetic color images with different spatio-colorimetric results according the weight given to spatial and color information.

Finally, we discuss previously obtained results, and propose further development taking into account the spatial information during the classification process, both in the decision and in the learning steps.

Compact histogram

Segmentation methods based on the analysis of color histograms are facing the difficulty of treating a huge quantity of information. For a color image of resolution $N \times M$ with each component coded on 8 bits, a standard 3D histogram is an array of 2^{24} cells, the number in each cell being coded on at least $\log_2(MN)$ bits in order to store the greatest number of pixels. In the case where $M=N=256$, the standard 3D histogram requires 128 Mo.

A few years ago, we proposed a new way of coding the *n*D histograms, leading to the so-called compact histogram [11]. Considering that most cells of the standard histogram are empty, the compact histogram retains only the C occupied cells. It consists of two arrays (figure 1): an array of size C×3 to store the colors, sorted out in lexicographical order, and an array of size C×1 for the corresponding populations of pixels. Since C is lower than MN, the compact histogram occupies less memory space, although it contains the full color information present in the image. For a 256×256 image with color components coded on 8 bits, the memory space required is less than 500 ko.

R	*G*	*B*
0	0	5
0	0	23
...	...	...
255	10	0
255	251	254

population
13
5
...
21
3

Figure 1. *Example of 3D compact histogram for a RGB color image (8 bits per component).*

Compact spatial neighborhood probability matrix

Taking into account previous researches in spatio-colorimetric classification such as [3] and [4], it is interesting to have a structure like a co-occurrence matrix, containing color pixels neighbors information.

For a given spatial direction, a co-occurrence matrix calculates how often a pixel with a color $c\alpha$ occurs adjacently to a pixel with a color $c\beta$. Other spatial relationships between pixels may be specified. That kind of structure is generally used to analyse grey levels textures. Without any requantization, a memory space of 256^2 cells is occupied.

A color image presents a maximum of 256^3 colors, its corresponding standard co-occurrence matrix will have 256^6 cells. Let be C the number of different colors in an image, C is usually lower than 256^3. Nevertheless, a co-occurrence matrix with C^2 cells requires a huge memory space. That is why a Spatial Neighborhood Probability Matrix (SNPM), requiring a reduced memory space, has been proposed.

This matrix specifies the probability to find a pixel with the color $c\beta$ in the neighborhood of a pixel with the color $c\alpha$, knowing that we have this pixel with the color $c\alpha$.

Given a color pixel $p_{c\alpha}$, the neighborhood $\boldsymbol{v}_d(p_{c\alpha})$ is the set of all pixels whose distance from $p_{c\alpha}$ is equal to d. For d=0, it is the point itself. For d=1, this neighborhood is defined by a 4 or 8 connexity. For d>1, a form (circle, square, diamond…) defines the connexity. The full neighborhood $\boldsymbol{V}(p_{c\alpha})$ is then defined as :

$$V(p_{c\alpha}) = \bigcup_{i=0}^{d} v_i(p_{c\alpha}) \tag{1}$$

Let be $\boldsymbol{h}_d(p_{c\alpha})$ the compact histogram associated with $\boldsymbol{v}_d(p_{c\alpha})$. A weighed histogram $H(p_{c\alpha})$ corresponding to $\boldsymbol{V}(p_{c\alpha})$ is defined as:

$$H(p_{c\alpha}) = \bigcup_{i=0}^{d} h_i(p_{c\alpha}) \times (1 + d - i) \tag{2}$$

where × is a weighting operator which multiplies each compact histogram $\boldsymbol{h}_i$ populations. Thus in $\boldsymbol{V}(p_{c\alpha})$, colors weights are higher when colors correspond to pixels close to $p_{c\alpha}$. If $\{c\alpha\}$ is the set of all pixels having the color $c\alpha$, $H_{c\alpha}$ is defined as:

$$H_{c\alpha} = \bigcup_{j=1}^{Card(\{c\alpha\})} H(p_{c\alpha_j}) \tag{3}$$

The SNPM is a cell array of size Cx1, with C the number of different colors in the image. Each cell *i*, $1 \leq i \leq C$, contains the corresponding $H_{c\alpha i}$ whose population has been normalized to 1 in order to express probabilities. The probability $P(c\beta \in H_{c\alpha})$ to find a pixel with the color $c\beta$ in the neighborhood of a pixel with the color $c\alpha$ in the image is directly given by the $c\beta$ entry of the histogram associated with the $c\alpha$ cell in the SNPM. By construction, SNPM contains C compact histograms, each histogram having less than C cells. Memory space required by SNPM is lower than the one required by a co-occurrence matrix, even coded in a compact form with C^2 cells.

The spatial distance $ds(c\alpha, c\beta)$ is defined from SNPM as the minimum between $[1-P(c\beta \in H_{c\alpha})]$ and $[1-P(c\alpha \in H_{c\beta})]$.

Hierarchical classification

The classification process through an unsupervised analysis of color histograms is an original 3D extension of the 2D limited method proposed in [9]. This method can be classed with watershed algorithm. Colors classification is carried out in two steps: the learning step and the decision step.

The learning step is a hierarchical decomposition of populations in the 3D histogram. For each level of population p_n, peaks P_i are identified by a connected components labelling process: first, the color compact histogram is thresholded for populations greater than or equal to p_n and a binary 3D matrix is reconstructed in the same way as a standard histogram but with only one bit per cell. Secondly, the binary matrix is labelled in 3D connected components. Each peak identified by a connected component is then iteratively decomposed into narrower peaks, beginning from population 0. Thanks to [10], the algorithm considers only existing populations in the compact histogram, jumping from one to the next in ascending order. A peak is then labelled as significant if it represents a population greater than or equal to a threshold S (expressed in percent of the total population in the histogram). The procedure is illustrated in figure 2 (drawn in one dimension for clarity). We shall name kernels K_i the peaks corresponding to circled leaves in part b of figure 2. In other words, kernels are significant peaks (part of figure 2) without descendants in the hierarchical decomposition tree (e.g., figure 2 shows five significant peaks P_i ($i = 0$ to 4) and three kernels K_i ($i = 2, 3, 4$)). The number of classes N_c is taken equal to the number of kernels (the class corresponding to kernel K_i is noted C_i). Therefore N_c depends on the threshold S, i.e. on the precision the image colors are analyzed with.

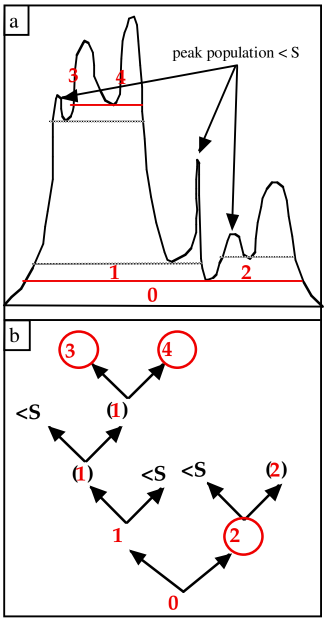


Figure 2. *An example of hierarchical decomposition. The circled leaves (part b) correspond to significant peaks as obtained at the end of the iterative decomposition (solid lines in part a), whereas leaves marked < S (part b) correspond to insignificant peaks (dotted lines in part a).*

In the decision step, the mass center $\mu(K_i)$ of each kernel K_i is calculated in the color space. Let us denote by $c\beta$ the color corresponding to the point of coordinates (r,g,b) in the color space. Two cases appear: if (r,g,b) belongs to K_i, color $c\beta$ is attributed to class C_i ; if not, let us denote by P_k the peak (r,g,b) it belongs to; color $c\beta$ is attributed to class C_i corresponding to kernel K_i, son of P_k, such that $d[\mu(K_i), (r,g,b)]$ is minimum, where $d[c\alpha, c\beta]$ is a spatio-colorimetric distance between $c\alpha$ and $c\beta$.

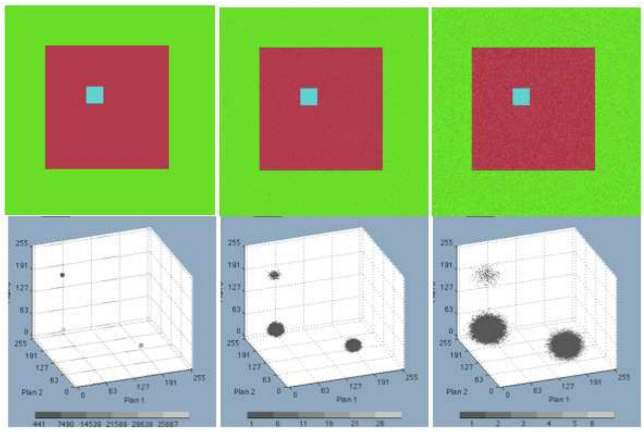
The spatio-colorimetric distance (dSC) is expressed by equation (4):

$$dSC(c\alpha, c\beta) = \theta\, ds(c\alpha, c\beta) + (1-\theta)\, dc(c\alpha, c\beta) \qquad (4)$$

where $dc(c\alpha, c\beta)$ is a normalized colorimetric distance between $c\alpha$ and $c\beta$ like Euclidean or Mahalanobis and $ds(c\alpha, c\beta)$ the spatial distance as defined from the Compact SNPM. θ, $0 \leq \theta \leq 1$, gives more or less weight to the spatial or colorimetric information.

Results

A synthetic RGB color image with 256x256 pixels, coded on 24 bits (8 bits per component) each, is used as a probe image to test the algorithm. This image is composed of three regions with pure colors: 2 regions with an important population the red and the green ones, and another composed of few pixels, the blue one. As shown in the RGB histogram (figure 3,a2), the blue region is colorimetrically closer to the green region than the red one.



(a1,a2) (b1,b2) (c1,c2)

Figure 3. *This figure is composed of three images: a1, b1 and c1, and their corresponding histogram a2, b2 and c2. The image a1 is the original probe color RGB image (8 bits per component). The images b1 and c1 are corrupted probe images; b1 is named lowNoise and c1 highNoise. In the histogram a2, the points have been enlarged for a better visibility.*

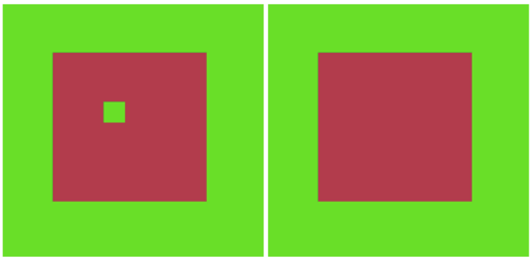
The RGB-histogram (a2) reveals the presence of three classes. In order to evaluate the interest of the classification algorithm, the colorimetric components of the probe image have been corrupted by an additional uncorrelated Gaussian noise: being N_{mr}, N_{mg}, N_{mb} three matrices of marginal centered noise with the same standard deviation $\sigma = 0.02$ for the image lowNoise (figure3 b1) and $\sigma = 0.05$ for the image highNoise (figure3 c1). The alteration of the probe image constituted of the three colorimetric planes P_r, P_g, P_b is given by $P_{iN}=P_i+N_{mi}$ ($i \in \{r,g,b\}$).

Due to a lower standard deviation, in lowNoise, some pixels from the blue region (the smallest one) have the same value, whereas in highNoise, all pixels in this region have a different value.

Considering that the blue region's population is insignificant, the threshold S (defined in the Hierarchical classification part) has been adjusted consequently. The goal of the segmentation is to find two classes, one for the red region and the other for the green one.

Firstly, the classification algorithm has been run with θ = 0 (defined in the Hierarchical classification part) to neglect the spatial information. The segmentation obtained is the same for the three images, and is presented in the figure 4(a). Two classes have been identified, one corresponding to the green region, another corresponding to the red region. Insignificant blue values have been classified with the green region that is colorimetrically the closest.

On another hand, θ has been adjusted to take into account the spatial information (θ > 0). The neighborhood distance d (defined in the Compact spatial neighborhood probability matrix part) has been chosen to cover the blue region while remaining inside the red region. Results are presented in the figure 4(b). The segmentation obtained is the same for the three images. Two classes have been obtained, one corresponding to the green region only, and the other one merging the blue and the red regions. Spatially, blue pixels are closer to red ones and so dSC (blue, red) is lower than dSC (blue, green).



(a) (b)

Figure 4. *This figure is composed of two images: a and b.*
(a) is the classification result of figure 3 (a1, b1 and c1) when $\theta = 0$.
(b) is the classification result of figure 3 (a1, b1 and c1) when $\theta > 0$, and with a value of neighborhood distance d covering the blue region.

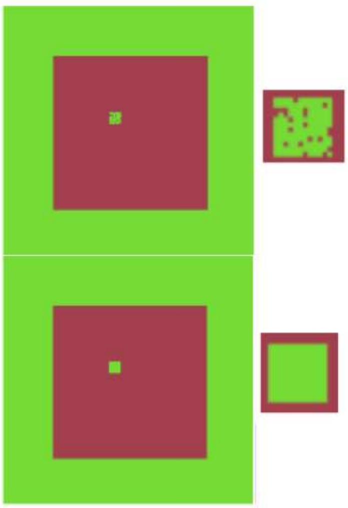
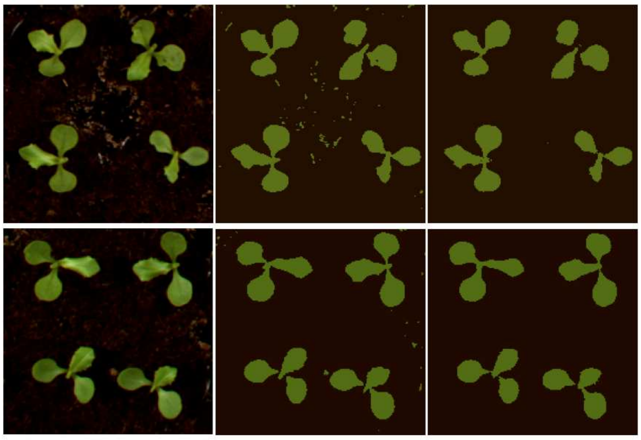


Figure 5. *This figure is composed of two parts, the upper part presents lowNoise's segmentation with a zoom on the original blue region. The lower part, presents the segmentation results for highNoise.*

Keeping θ > 0, but reducing d so that the neighborhood becomes smaller than the blue region, the classification result for the probe image remains the same because dSC (blue, red) remains lower than dSC (blue, green). In noisy images, the result is different (figure 5): only edges of the blue region have been classified with red region. As in lowNoise a few pixels from the centre of the blue region have the same colorimetric value as pixels from the edge, they are also classified with the red region. In highNoise, all the blue pixels have different colorimetric values, so only edges become red, and the others pixels are merged with green ones which are colorimetrically closer.

The method has been applied to real images of plants (lettuces) in order to extract plants from loam, taking in account the heterogeneity of the loam. In the first column (a) of Figure 6, two real images are illustrated, they have been photographied with a tri-CCD camera, and lettuces were highlighted with a halogen lamp. In the second column (b), color hierarchical classification results are presented, θ was equal to 0, in order not to take care of the spatial information. In the third column (c), the spatio-colorimetric results with θ=0.5 are shown.



(a) (b) (c)

Figure 6. *This figure is composed of six images, in three columns a, b and c.*
(a) real RGB images of 9-day lettuce plantlets from two different variety.
(b) color classification ($\theta = 0$).
(c) spatio-colorimetric classification, with θ=0.5 and neighborhood distance d=4.

With hierarchical color only classification, which is a watershed like process, some bright and very small regions of loam are recognized as plants. The spatio-colorimetric classification is able to classify correctly these regions as loam. In these experiments, it was important to limit the use of spatial filtering taking in account the area of the regions because some slow growing lettuces are very small and would be eliminated with filtering.

Discussion and perspectives

Significant results have been obtained with the proposed unsupervised vectorial hierarchical spatio-colorimetric classification. A large amount of plants images were segmented automatically. The comparison with experts segmentations is in progress, and first results are very encouraging.

However, the method relies on two new parameters: θ and the neighborhood distance d, which are difficult to fix without an *a priori* knowledge of the image to be segmented. The results obtained have shown efficiency for a neighborhood covering small regions, with insignificant population. These regions cannot correspond to classes during the learning step regarding threshold S. They are merged during the second step, where the spatial information has been introduced. The weight of the spatial information depends on θ, which is correlated to the colorimetric distance between these small regions and the others, that is why both θ and d are difficult to evaluate *a priori*.

Wider colorimetric regions could be treated introducing the spatial information during the learning step of the classification. Actually, if a color population is high enough, it will be a kernel in the histogram and will form a class. If this class does not correspond to a spatial region, the problem cannot be solved increasing the neighborhood distance d. On the other hand, the hierarchical decomposition of the histogram could be constrained to form classes which satisfy a spatio-colorimetric homogeneity criterion.

References

[1] Trémeau A., Laget B., "Quantification de la couleur et analyse d'image", Traitement du signal. (1995).

[2] Xuan G., Fisher P.. Maximum likelihood clustering method based on color features. Proceedings of the First International Conference on Color in Graphics and Image Processing, Saint-Etienne, France, p. 191-194 (2000).

[3] Trémeau A., Contribution des modèles de la perception visuelle à l'analyse d'image couleur, France, PhD thesis, University of Saint-Etienne, (1993).

[4] Macaire L., Vandenbroucke N., Postaire J.G., "Color image segmentation by analysis of subset connectedness and color homogeneity properties", Computer Vision and Image Understanding, Elsevier, 102:105–116, (2006).

[5] Foucher P., Revollon P., Vigouroux B., "Segmentation d'images en couleurs par réseau de neurones : application au domaine végétal", Actes du congrès francophone par vision par ordinateur (ORASIS), Cahors, France, 309-317, (2001).

[6] Busin L., Vandenbroucke N., Macaire L., Postaire J.G., "Colour space selection for unsupervised colour image segmentation by analysis of connectedness properties", International Journal of Robotics and Automation, 20(2):70-77, (2005).

[7] Noordam J., Broek W.V.D., "Geometrically guided fuzzy c-means clustering for multivariate image segmentation", In International Conference on Pattern Recognition, 1:462–465, (2000).

[8] Comaniciu D., Meer P., "Mean Shift: A Robust Approach toward Feature Space Analysis", IEEE Trans. Pattern Analysis Machine Intell., 24(5):603-619, (2002).

[9] Clément A., Vigouroux B., Unsupervised segmentation of scenes containing vegetation (Forsythia) and soil by hierarchical analysis of bi-dimensional histograms. Pattern Recognition Letters, n° 24, p. 1951-1957, (2003).

[10] Ouattara S., Clément A.. Unsupervised Image Segmentation by Multi-Dimensional Compact Histograms Analysis. ClasSpec'08, October 15th 2008, Lens, France, (2008).

[11] Clément A., Vigouroux B., Un histogramme compact pour l'analyse d'images multi-composantes. Actes du 18ème colloque sur le traitement du signal et des images GRETSI' 01, Toulouse, France, vol. 1, p. 305-307, (2001).

Author Biography

Cindy Torres is a Phd student in image processing at the Laboratoire d'Ingénierie des Systèmes Automatisés (LISA) at the University of Angers, France. In 2007, she obtained her French diploma in "BioSciences Engineering" from the Institut Supérieur des Biosciences de Paris (ISBS). Her research interests include analysis, classification and segmentation of color images. She is although implicated in automated plants phenotyping for industry by means of image processing.

Alain Clément received his PhD degree in Image and Signal Processing from the University of Angers, France, in 2002. He is currently a senior lecturer in the University Institute of Technology of Angers where he teaches Computer Science, Networking, Data Bases and Image Processing. His research interests at the LISA laboratory are focused on color and multicomponent image processing, especially on classification methods by analysis of multidimensional histograms.