Scale space filter based on homogeneity degree for color image segmentation

L. Busin^(1,2), N. Vandenbroucke^(1,2), L. Macaire⁽¹⁾ and J.-G. Postaire⁽¹⁾ laurent.busin@eipc.fr, nicolas.vandenbroucke@eipc.fr, ludovic.macaire@univ-lille1.fr, jack-gerard.postaire@univ-lille1.fr

⁽¹⁾Laboratoire LAGIS UMR CNRS 8146 Université des Sciences et Technologies de Lille Bâtiment P2 - 59655 Villeneuve d'Ascq - FRANCE ⁽²⁾Ecole d'Ingénieurs de Pas-de-Calais Campus de la malassise 62967 Longuenesse - FRANCE

Abstract

Scale space filtering is a well known approach in image segmentation. This approach determines precisely the thresholds in order to produce the image segmentation by histogram multi-thresholding. The main drawback of this approach is that it produces a coarse segmentation because it does not take into account the spatial arrangement of the pixels in the image and so, requires a second stage in order to obtain a finer segmentation. In this paper, we propose to associate a new criterion, the compactness degree, with the scale space filter in order to produce the segmentation by means of one single stage. For this purpose, this criterion exploits the connectedness and the homogeneity properties of pixels.

Introduction

The color image segmentation techniques described in the literature can be categorized into two main classes, depending on the distribution of the pixel colors is analyzed either in the image plane or in a color space [1]. In this paper, we propose an approach which takes into account both connectedness properties of the pixels in the image plane and homogeneity properties of the pixels in the color space.

Among the techniques analyzing the pixel colors in a color space, color image segmentation by 1D-histogram multithresholding assumes that homogeneous regions in the image plane give rise to peaks of the one-dimensional histogram of each color component of a color space. 1D-histogram multithresholding consists in determining the thresholds delimiting these peaks in order to construct classes of pixels in the threedimensional color space and produce the image segmentation.

Many authors apply the *scale space filter* in order to determine these thresholds [2, 3]. The scale space filter partitions the histogram of each color component into intervals containing only peaks and valleys. Since this partition depends on specific criteria which do not take into account the spatial arrangement of the pixels in the image, it produces a coarse segmentation of the color image. Therefore, it is necessary to perform a spatial analysis of the coarsely segmented image in order to obtain a fine segmentation [4].

In this paper, we propose a new criterion associate with the scale space filter to select the most significant peaks of a 1D-histogram. Since this criterion takes simultaneously into account the spatial arrangement of pixels in the image and the dispersion of their levels, the color image can be segmented by one single stage.

In the second section, we present the detection of peaks of the 1D-histogram by the scale space filter. Then, we detail the *compactness degree* which is the criterion used to select the thresholds delimiting the most significant peaks. In the fourth section, we detail the color image segmentation based on pixel classification deduced from the so-determined thresholds. Experiments on a synthetic color image and natural color images are carried out in the last section.

Scale space filter

In the first part of this section, we introduce the scale space filter. Then, we show how this filter can be applied to segment a synthetic color image for the illustration purposes.

Principle

The scale space filter precisely detects the most significant peaks and valleys of an histogram [5]. For this purpose, the histogram is smoothed with different Gaussian functions. This smoothed histogram is the result of a convolution between f(x), the signal corresponding to the considered histogram (where x is the color component level) and $g(x, \tau)$ a Gaussian kernel (where τ is the standard deviation). The convolution, denoted '*', between these two functions defines the function $F(x, \tau)$ which represents the smoothed histogram:

$$F(x,\tau) = \int_{-\infty}^{+\infty} f(u) \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{(x-u)^2}{2\tau^2}\right] du \tag{1}$$

The analysis of the derivatives of an histogram allows to determine the extrema which locate the peaks and the valleys. So, for different values of the *scale* τ , the zero-crossings of the derivatives of the smoothing histogram are detected and represented in the scale space (x, τ) in order to construct the *finger*-*prints*. Here, we use the zero-crossings of the first derivative $\partial F/\partial x$ in order to determine local minima and local maxima of the smoothed histogram. The local maxima define the centers of the peaks and the local minima correspond to the centers of valleys which are also the thresholds delimiting the peaks.

The fingerprints are analyzed in order to obtain an *interval tree* of the scale space (x, τ) which represents all the possible partitions of the histogram into peaks and valleys. Each interval in the interval tree is called a *node*. In this paper, only peaks are detected. So, each node in the scale space corresponds to a peak of the histogram and is defined by a rectangle which is delimited by a left threshold, a right threshold, a down scale and an up scale. Let T_x^k , be the right threshold of the k^{th} node for a given scale τ along the *x* color component. Thus, T_x^{k+1} is the right threshold of the $(k+1)^{\text{th}}$ node along the *x* color component but also the left threshold of the k^{th} node.

Example

In order to illustrate the scale space filter, we use the synthetic color image of figure 1(a) which is composed of six regions:

- a brown background,
- a yellow little square,
- an orange big square,
- a purple patch,
- two concentric green disks.

More precisely, we apply the scale space filter to the blue component of this synthetic color image where the pixels are only characterized by the blue levels.



(a) Synthetic color image(b) Blue component imageFigure 1. The synthetic color image and the blue component image.

Figure 2 represents the histogram of the blue component image of figure 1(b).



Figure 2. Histogram of the blue component.

This histogram shows three peaks which corresponding to three pixel classes :

- the class of background and two squares pixels,
- the class of interior green disk pixels,
- the class of exterior green disk and purple patch pixels.

By smoothing the histogram of figure 2 with respect to different increasing values of τ (see equation (1)), we can detect the local minima and the local maxima of the smoothed histogram at different scales. The locations of these extrema at different scales define chains of adjacent points which are the fingerprints in the scale space. Figure 3 shows the fingerprints provided by the multi-scale analysis of the blue component image of figure 1(b).



Figure 3. Fingerprints corresponding to the blue histogram.

In this figure, x is the blue level of the image of figure 1(a) and τ is the scale. The fingerprints labeled as green represent the local maxima at different scales of the scale space, that is to say the centers of the peaks in the histogram and the fingerprints labeled as red represent the local minima at different scales of the scale space, that is to say the centers of the valleys in the histogram. Here, we analyze only the red fingerprints which delimits the peaks of the histogram.

Then, by tracking each fingerprint from the top of the scale space to the bottom, it is possible to determine the nodes. Each node is define by a rectangle which is delimited by four sides :

- an up scale which is the scale from that, a new fingerprint starts,
- a down scale which is the scale from that a fingerprint stops and/or new fingerprints start,
- left and right thresholds which correspond to the locations of the local minima for the up scale and the down scale.

In this figure, the nodes are represented in blue. When τ is too high, the smoothed histogram presents only one peak which is delimited by the interval [0;255]. The rectangle of the corresponding node is then delimited by the values of this interval, a down scale and an up scale which is defined as the value of the down scale plus one. Figure 4 represents the nodes deduced from the analysis of the fingerprints of figure 3.



Figure 4. Determination of the nodes from the fingerprints.

Finally, the determined nodes define an interval tree in the scale space which is represented in figure 5.



Figure 5. Interval tree.

This figure shows that 7 nodes are reconstructed. For $\tau \leq 1.15$, four rectangles or nodes are reconstructed meaning that four peaks are detected. For $1.15 < \tau \leq 3.95$, three peaks are detected. For $3.95 < \tau \leq 24.2$, two peaks are detected. For $\tau > 24.2$, only one peak is detected.

Then, the nodes corresponding to the most significant peaks must be selected thanks to a specific criterion. These selected nodes are called *active nodes* and are used to produce the segmentation.

Authors generally consider that a node is active when its height (the difference between its up scale and down scale) is higher than the mean of its off-springs because it is the most stable in the scale space [6]. Since this criterion does not take into account the spatial arrangement of the corresponding pixels in the image, its use produces a coarse segmentation. So, we introduce a new criterion, the *compactness degree* which simultaneously measures the spatial arrangement of the pixels and the dispersion of their levels [7].

Compactness degree

The compactness degree is defined as the product of a *connectedness degree* and a *homogeneity degree*. The connectedness degree of a subset of pixels is a measure of the spatial arrangement of its pixels in the image plane. The homogeneity degree of a subset of pixels reflects the dispersion of the levels representing its pixels.

Let us denote the subset $S[T_x^k, T_x^{k+1}]$ of pixels whose color component levels x range between T_x^k and T_x^{k+1} in the considered image, that is to say which belong to the detected peak delimited by T_x^k and T_x^{k+1} . For the sake of simplicity, the subset $S[T_x^k, T_x^{k+1}]$ will be denoted hereafter S.

Connectedness degree

Let $N_S(P)$ be the subset of the 8 neighboring pixels of P which belong to S. The connectedness between P and the subset S, denoted $\gamma_S(P)$, depends on the number of pixels which belong to $N_S(P)$. It is defined as:

$$\gamma_{\rm S}(P) = \frac{\operatorname{Card}\{N_{\rm S}(P)\}}{8},\tag{2}$$

where the number 8 is a normalizing factor.

In order to define a connectedness measure of a nonempty subset of pixels, we introduce the connectedness degree of a subset *S*, denoted CD(S), which is defined as:

$$CD(S) = \frac{\sum\limits_{P \in S} \gamma_S(P)}{\operatorname{Card}\{S\}}.$$
(3)

The connectedness degree of an empty subset of pixels is set to 0. The connectedness degree CD(S) is the mean number, normalized by 8, of neighbors of the pixels of *S* which belong also to *S*. A low connectedness degree of a subset *S*, close to 0, means that its pixels are sparsely scattered through the image, while a high connectedness degree, close to 1, indicates that its pixels are strongly connected in the image.

When a interval $[T_x^k, T_x^{k+1}]$ is too large, it may contain different well separated peaks in the histogram which correspond to different regions of different colors. The analysis of the connectedness degree is not sufficient for discriminating a subset which corresponds to an actual region in the image from a subset which corresponds to several regions with different colors. Hence, the segmentation procedure has also to take into account the level homogeneity properties of the subsets.

Homogeneity degree

So, we propose to use a new estimation of the homogeneity properties of a subset *S* that is based on a measure of the dispersion of the levels representing its pixels. This measure, denoted $\sigma(S)$, is estimated as:

$$\sigma(S) = \frac{1}{\operatorname{Card}\{S\}} \cdot \sqrt{\sum_{P \in S} (x(P) - M(S))^2},\tag{4}$$

where x(P) is the color component level of the pixel *P* and M(S) is the mean level of the pixels which belong to *S*.

In order to determine if a subset corresponds to an actual region, we propose to compare a global measure of the level dispersion of this subset with a local measure of the level dispersion at each pixel which belongs to the subset. For each pixel *P* of *S*, we determine the dispersion measure $\sigma(N_S(P))$ of the subset $N_S(P)$ constituted of the 8 neighbors of *P* that belong to *S*. Let $\sigma_{local}(S)$ be the *local dispersion measure* of the subset *S*, defined as the mean of the dispersion measures $\sigma(N_S(P))$ estimated for the subsets $N_S(P)$ of all the pixels *P* of *S*. This local dispersion measure of the subset *S* is expressed as:

$$\sigma_{local}(S) = \frac{1}{\operatorname{Card}\{S\}} \cdot \sum_{P \in S} \sigma(N_S(P)).$$
(5)

When the local dispersion measure $\sigma_{local}(S)$ of the subset *S* tends to be close to its dispersion measure $\sigma(S)$, there is a great probability that the levels of the pixels of the subset *S* give rise to a single peak, whereas when $\sigma_{local}(S)$ tends to be lower than $\sigma(S)$, the levels could be split into several well separated peaks along the 1D color-component histogram.

In order to compare $\sigma_{local}(S)$ and $\sigma(S)$, we define the homogeneity degree HD(S) of the subset *S* as:

$$\begin{cases} HD(S) = \frac{\sigma_{local}(S)}{\sigma(S)} & \text{if } \sigma(S) \neq 0, \\ HD(S) = 1, & \text{otherwise.} \end{cases}$$
(6)

The homogeneity degree HD(S) ranges from 0 when $\sigma_{local}(S)$ is equal to 0, to 1 when $\sigma_{local}(S)$ is equal to $\sigma(S)$. In order to avoid values of HD(S) higher than 1, HD(S) is set to 1 when $\sigma_{local}(S)$ is higher than $\sigma(S)$. When the levels of the pixels of the considered subset give rise to one compact peak along the color component, the homogeneity degree of the subset *S* is close to 1. On the other hand, when the levels of the pixels of the subset give rise to well separated peaks along the color component, the homogeneity degree of the subset *S* is low and close to 0.

Compactness degree

In order to combine connectedness and homogeneity concepts, we define the compactness degree of a subset S, denoted CHD(S), as the product of its connectedness degree and its homogeneity degree:

$$CHD(S) = CD(S) \cdot HD(S). \tag{7}$$

This compactness degree reaches its highest value 1 if the connectedness degree and the homogeneity degree are both equal to 1. A high compactness degree indicates that the pixels of the subset are strongly connected in the image (connectedness degree close to 1) and that the levels representing its pixels give rise to one compact peak in the histogram (homogeneity degree close to 1). On the other hand, the spatial-color compactness degree of a subset is close to 0 when either its connectedness degree of a subset means that its pixels are sparsely scattered through the image or that the levels representing its pixels do not give rise to a single compact peak in the histogram.

The compactness degree of each subset $S[T_x^k, T_x^{k+1}]$ of pixels belonging to each interval $[T_x^k, T_x^{k+1}]$ of the interval tree in the scale space is then estimated. For example, the values of the compactness degree for different partitions of the histogram of figure 2 are given in table 1.

Table 1 shows that, when the scale τ decreases the number N of intervals increases because a *parent interval* in the interval tree gives rise to two *child intervals* at least for a specific value of τ in the scale space. So, when two child intervals are created, theirs compactness degrees are compared with those of the parent interval. If the compactness degree of the child intervals are all higher than the compactness degree of the parent

Number of	Interval of τ			
Intervals	Interval $[T_B^k, T_B^{k+1}]$			
N	Compactness degree $CHD(S[T_B^k, T_B^{k+1}])$			
	$\tau > 24.2$			
1	[0;255]			
	0.197			
	$3.95 < au \leq 24.2$			
2	[0;72]		[72;255]	
	0.689		0.526	
	$1.15 < \tau \le 3.95$			
3	[0;72]		[72;106]	[106;255]
	0.689		0.719	0.793
	$\tau \le 1.15$			
4	[0;72]	[72;81]	[81;106]	[106;255]
	0.689	0	0.720	0.793



interval, the corresponding node can't be selected as an active node. The active nodes are selected as those from whose the value of the compactness degree does not decrease. For example, the value of the compactness degree for the interval [72;255] (N = 2) is 0.526. This interval gives rise to the intervals [72;106] and [106;255] (N = 3) for which the values of the compactness degree are 0.719 and 0.793 respectively. These values are both higher than 0.526. Here, the child intervals correspond to pixels more connected in the image and to peaks more compact in the histogram than those corresponding to the parent interval. So, the node corresponding to the interval [72;255] can't be selected as an active node. In this example, the active nodes are those which correspond to N = 3 because the value of the compactness degree corresponding to the child interval [72;81] (N = 4) is lower than those corresponding to the parent interval [0;72] (N = 3).

Image segmentation

In order to segment a color image by using the thresholds of the three 1D color components, it is necessary to determine the classes of pixels in the 3D color space. In this space, a class of pixels can be defined by a parallelepipedal box. This box is delimited by thresholds of the selected active nodes which partition each of the three 1D color histograms. Figure 6 shows the partition of the color space into $N_R \times N_G \times N_B$ parallelepipedal boxes where N_R , N_G and N_B are the number of active nodes along the red, green and blue components respectively.



Figure 6. Partition of the (R,G,B) color space into parallelepipedal boxes.

Thus, this partition defines $N_R \times N_G \times N_B$ possible classes

of pixels. For the image segmentation by pixel classification, it is necessary to select only valid classes from this partition. The valid classes are defined by the boxes in which fall into the color vectors with the highest populations and which correspond to compact subsets of pixels. The pixels whose color vectors fall into these boxes constitute the N_C classes of pixels.

Experimental results

In order to demonstrate the interest of our proposed approach, we detail the results obtained with the synthetic color image of figure 1(a) in the first part of this section. Then, we present different results obtained with natural benchmark color images commonly used to compare the relevance of image segmentation methods.

Synthetic image result

We propose to compare the segmentation result of the image of figure 1(a) performed by the selection of the active nodes using the classic criterion (the height) and those obtained when using the compactness degree. Image 1(a) is constituted by six distinct regions while the histogram of the blue component of this image contains three significant peaks (see figure 2).

Figure 7(a) shows that only two active nodes, and so two significant peaks are detected by using the classical criterion applied to the histogram of figure 2. For this case, we obtain $N_R = 2$, $N_G = 2$ and $N_B = 2$. So, there are only $N_R \times N_G \times N_B = 8$ possible classes. Figure 7(b) shows the labeled pixels classified to the four classes constructed thanks to the classification procedure. For the illustration purposes, the label of each class does not coincide with its color center. The color segmentation does not succeed in separating the regions corresponding to the two concentric green disks and the two squares (see image 7(b)). The number of valid classes is only $N_C = 4$.



Figure 7. Results obtained by the scale space filter using the classical criterion.

Table 1 shows that the value of the compactness degrees stop to increase when four actives nodes are considered. So, three active nodes must be selected. For this case, we obtain $N_R = 2$, $N_G = 4$ and $N_B = 3$. So, there are $N_R \times N_G \times N_B = 16$ possible classes. Figure 8(a) shows the three active nodes of the blue image determined by the scale space filter integrating the compactness degree. Figure 8(b) shows that the six classes of pixels ($N_C = 6$) are well constructed thanks to our segmentation scheme. Only a few pixels representing the two green concentric discs are misclassified since the clusters of their color points strongly overlap in the (R, G, B) color space.



Figure 8. Results obtained by the scale space filter using the analysis of the compactness degree.

Natural image results

In this part, we show some results obtained by our approach applied to these three benchmark images:

- "Jelly beans" image (see figure 9(a)): this image represents four kinds of colored objects on an uniform background.
- "Hand" image (see figure 10(a)): this image contains a hand on a textured background and a blue ring around a finger.
- "House" image (see figure 11(a)): six main classes of pixels can be determined in this image, namely the sky, the wall, the window, the shadows, the roof and the gutters.

By applying our approach to these images, we obtain the results of figures 9(b), 10(b) and 11(b). The pixels assigned to the same class are labeled with a false color corresponding to this class.

The obtained results on the "Jelly beans" image are remarkable because the five classes of pixels are precisely detected. For this image, we obtain $N_R = 2$, $N_G = 2$, $N_B = 2$ and $N_C = 5$.



(a) Original image (b) Lab Figure 9. "Jelly beans" image.

(b) Labeled image by our approach

In the "hand" image, three classes of pixels are principally reconstructed: the hand, the background though it is composed of textured colors and the ring. Nevertheless, all the pixels representing the ring are not assigned to the same class because the ring is compounded by two main colors. A fourth class of few pixels corresponding to highlights is also reconstructed. For this image, we obtain $N_R = 2$, $N_G = 1$, $N_B = 2$ and $N_C = 4$.

The obtained results on the "Hand" image are satisfying because all the classes of pixels are correctly reconstructed. Few pixels are not assigned to any class because their colors fall into boxes of the color space which are not considered as classes. These pixels mainly correspond to edges of objects. For this image, we obtain $N_R = 3$, $N_G = 3$, $N_B = 2$ and $N_C = 6$.





(b) Labeled image by our approach

(a) Original image *Figure 10. "Hand" image.*





(a) Original image *Figure 11. "House" image.*

(b) Labeled image by our approach

Conclusion

The first results obtained with the proposed approach in this paper are very encouraging. Indeed the tested images are correctly segmented while only the (R, G, B) space is used. It would be interesting to apply our approach with different color spaces to study its behavior and more precisely the relationship between the compactness degree and the color space. Presently, we work on the automatic selection of the most adapted color space for a problem of image segmentation [8] which exploits the scale space filter based on the compactness degree.

Presently, the value of the compactness degree is tracked scale by scale in the interval tree. But, it would be interesting to analyze the compactness degree on each branch of the interval tree independently and to study the evolution of this criterion in order to select the number of active nodes.

Another improving point is to integrate this approach to an iterative segmentation algorithm in order to extract pixel classes step by step and to determine the most adapted color space to construct each class at each step [9].

References

- H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 2259–2281, 2001.
- [2] 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.
- [3] H. D. Cheng and J. Li, "Fuzzy homogeneity and scale-space approach to color image segmentation," *Pattern Recognition*, vol. 36, no. 7, pp. 1545–1562, 2003.
- [4] C. L. Huang, T. Y. Cheng, and C. C. Chen, "Color images' segmentation using scale space filter and markov random field," *Pattern Recognition*, vol. 25, no. 10, pp. 1217–1229, 1992.
- [5] R. Schettini, "A segmentation algorithm for color images," *Pattern Recognition Letters*, vol. 14, no. 8, pp. 499–506, 1993.
- [6] A. P. Witkin, "Scale-space filtering: a new approach to multi-scale description," *Image Understanding*, pp. 79–95, 1984.
- [7] L. Macaire, N. Vandenbroucke, and J.-G. Postaire, "Color image

segmentation by analysis of subset connectedness and color homogeneity properties," *Computer Vision and Image Understanding*, vol. 102, no. 1, pp. 105–116, 2006.

- [8] L. Busin, N. Vandenbroucke, L. Macaire, and J.-G. Postaire, "Colour space selection for unsupervised colour image segmentation by analysis of connectedness properties," *International Journal of Robotics and Automation*, vol. 20, no. 2, pp. 70–77, 2005.
- [9] L. Busin, N. Vandenbroucke, L. Macaire, and J. G. Postaire, "Color space selection for unsupervised color image segmentation by histogram multithresholding," in *Proc. of the IEEE Int. Conf. on Image Processing*, Singapour, 2004, pp. 203–206.

Author Biography

Laurent Busin, Engineer (Ecole d'Ingénieurs du Pas-de-Calais) is a PhD student at the Institute of Technology of the Université des Sciences et Technologies de Lille, France. His research concerns color image segmentation and quality control for an industrial application. This industrial application consists to analyse the color pattern glasses.

Nicolas Vandenbroucke received the PhD in computer science and control from the University of Lille in 2000. Since 2002, he is an associate professor in the Ecole d'Ingénieurs du Pas-de-Calais. He is a member of the LAGIS research laboratory in the "Vision and Image" group. His research interests include color representation, color image analysis applied to segmentation.

Ludovic Macaire received the M.S (Engineer) degree in Computer Science from the UTC Engineering school of Compiegne, France, in 1988; the PhD in computer science and control from the University of Lille in 1992, and the Habilitation à Diriger des Recherches from the University of Lille in 2004. Since 1993, he is an associate professor in the LAGIS Laboratory (Laboratoire d'Automatique, Génie Informatique & Signal) at the University of Lille. His research interests include color representation, color image analysis applied to segmentation and retrieval.

Jack-Gérard Postaire, Engineer (Ecole Centrale de Lille, France, 1970), Docteur es Sciences (U.S.T.L., Lille, France, 1981) is a Professor in Automatic Control and Computer Science at the Institute of Technology of the Université des Sciences et Technologies of Lille, France. He has been at the U.S.T.L. since 1982, when he was promoted to the rank of Professor at the age of 36. Professor Jack-Gérard Postaire has authored or co-authored over 200 scientific conferences and more than 60 scientific papers. He has authored one book on Pattern Classification and Image Analysis. Professor Jack-Gérard Postaire has directed over 40 research projects, yielding several million euros of funding and fees. He has also supervised over 50 thesis and is a regular reviewer for many international scientific journals. He is presently the head of the " Vision and Image " research group at the LAGIS (Laboratoire d'Automatique, Génie Informatique et Signal) UMR CNRS 8146, of the Université des Sciences et Technologies de Lille.