PERFORMANCE EVALUATION OF IMAGE SEGMENTATION. APPLICATION TO PARAMETERS FITTING

S. Chabrier, H. Laurent and B. Emile

Laboratoire Vision et Robotique - UPRES EA 2078
ENSI de Bourges - Université d'Orléans
10 boulevard Lahitolle, 18020 Bourges Cedex, France
phone: +33 2248484000, fax: +33 2248484040
email:sebastien.chabrier@ensi-bourges.fr

ABSTRACT

This paper deals with the performance evaluation of image segmentation. The goal of this work is to show some techniques that enable the comparison of different segmentation results. We first present a visualization method that facilitates the visual evaluation of a segmentation result. Next, we focus on unsupervised evaluation criteria that do not take into account any *a priori* knowledge to quantify the quality of a segmentation result. Finally, we use these evaluation criteria to determine the best fitted parameters of a segmentation method for a given image and a desired level of precision.

1. INTRODUCTION

Segmentation is one of the first step in an image analysis process and is of crucial importance. The goal is there to decompose the image in homogeneous regions in order to facilitate the scene interpretation which is done afterwards. That is why many segmentation methods have been proposed for the last years [1], [2], [8] ... The multitude of the available information in this domain is at once very attractive but also destabilizing for the user who is often placed in a tricky position in front of this prolific literature. It indeed remains difficult to evaluate the segmentation methods efficiency and to fix ones's choice on a single method, no one being optimal in all cases. We present in this communication two approaches that allow to evaluate the quality of a segmentation result and that can make the choice easier for an user.

In the first part of this article, we present a method that facilitates the visual evaluation of a segmentation result by using a colored representation. The classical presentation of different segmentation results in the grey-level domain often makes it difficult to compare their respective performances. The human eye is indeed incapable to distinguish between close grey levels. The change to the color domain mainly allows us to overcome this human limitation.

In a second part, we focus on unsupervised evaluation criteria allowing to quantify the quality of a segmentation result by considering different statistics without any *a priori* knowledge. Lots of evaluation criteria have been proposed in the literature to quantify the quality of a segmentation result [5], [7], [11] ... These criteria can be applied to many applications such as the comparison of segmentation methods or the choice of the best fitted parameters of a segmentation method for a given image. We compared in [3] six unsupervised criteria on uniform, textured and mixed, synthetic and real images. This study showed that two criteria seem to be the most relevant ones. The first one is favorably applied to textured images [9] and the other one is more adapted for uniform and low textured images [10]. In this article, we illustrate in a real case concerning aerial images processing, the interest of the two, above mentionned, unsupervised evaluation criteria for segmentation method grading.

Finally, some conclusions and perspectives are given.

2. VISUAL EVALUATION OF A SEGMENTATION RESULT

After having segmented a gray-level image by using different methods or a single one with different parameters, one has generally to determine which is the most satisfactory result. If some evaluation criteria can guide the user in his decision, it can be comfortable to be able to visually assess the results. We address and illustrate here the difficulty of the visual comparison of different segmentation results of an image. The EDISON algorithm [4] and a classical fuzzy C-means algorithm (FCM) [6] are used for this example.

One usual method to visualize a segmentation result is to attribute to each detected region the gray level corresponding to the average value of all pixels composing the region in the original image. As shown in figure 1, this method allows a good visual representation of a segmentation result. Nevertheless, when two adjacent regions have similar average value, it can become difficult to evaluate the quality of the segmentation result. The human eye is indeed not able to distinguish two regions with very close gray levels, the different regions can then seem to form a whole.

In order to overcome this problem, it is also possible to use a second colored image to display the segmentation result. Each region is represented with a random color chosen in a colormap uniformly spread among the RGB one. Let N_R be the number of regions of the image to be colored, the N_R color palette is first created. Each color is then randomly attributed to one region. This second image allows, as shown in figure 2, to clearly distinguish the borders of the regions.

Nevertheless, the same segmentation result presented with two random palettes can appear very different (see figure 3). If the proposed procedure is independently applied to two segmentation results, it will be, in that case, difficult to compare them.

To solve this problem, we propose a color matching procedure to make the visual assessment easier. We first apply the above mentioned coloring procedure to the segmentation result composed of the highest number of regions. We then obtain the reference colormap and the colored representation

The authors would like to thank financial support provided by the Conseil Régional du Centre and the European union (FSE).

(a) Original

(b) Edison (c) FCM

Figure 1: **Mean-gray level representations :** (a) Original image - (b) Segmentation by the Edison algorithm - (c) Segmentation by fuzzy C-means (FCM)

(a) Edison (b) Edison

Figure 2: **Mean-gray level and colored representations:** (a) Gray-level result (Edison segmentation) - (b) Colored result (Edison segmentation)

: I_{ref}. For each segmentation result left I_k, we search the region R of I_k having the highest intersection with a region X of I_{ref} and color this region R with the color of X. This color is then declared unusable and the process is repeated until no region of I_k has common pixels with one left region of I_{ref}. Finally, if some regions of I_k remain unpaired, they are randomly colored, taking one color among the I_{ref} unused ones. Figure 4 presents the colored visualizations corresponding to the segmentation results of figure 1. It becomes, in that case, much easier to visually pair the regions and to compare the quality of the different segmentation results.

(a) Edison (b) Edison

Figure 3: **Random coloring :** (a) First colored result (Edison segmentation) - (b) Second colored result (Edison segmentation)

(a) Edison (b) FCM

Figure 4: **Matched coloring :** (a) Colored result (FCM segmentation) - (b) Matched colored result (Edison segmentation)

3. UNSUPERVISED EVALUATION CRITERIA

We focus in this part on unsupervised evaluation criteria that enable to quantify the quality of a segmentation result without any *a priori* knowledge. Six unsupervised evaluation criteria of segmentation results had been compared in a previous work [3] (Zeboudj, Inter, Intra, Intra-inter, Borsotti and Rosenberger). A database including 300 synthetic images composed of textured and uniform regions was used for the comparison. This study revealed that two criteria seem to give better results than others: Zeboudj [10] and Rosenberger [9] criteria. Each criterion is more adapted for a type of image. Zeboudj's contrast is more adapted for uniform images, while Rosenberger's criterion is more relevant to textured images. We illustrate in this paper the interest of these two criteria for the fitting of parameters of a segmentation method.

- Zeboudj's contrast (Zeboudj) : [10]
  This contrast takes into account interior contrast and external contrast of the regions in the neighborhood of each pixel. If we note $W(s)$ a neighborhood of a pixel s, $f(s)$

the pixel intensity and L the maximum intensity, the contrast inside (I_i) and with outside (E_i) of the region R_i are respectively :

$$I_i = \frac{1}{A_i} \sum_{s \in R_i} \max \{c(s,t), t \in W(s) \cap R_i\} \qquad (1)$$

$$E_i = \frac{1}{l_i} \sum_{s \in F_i} \max \{c(s,t), t \in W(s), t \notin R_i\} \qquad (2)$$

where A_i is the surface and F_i is the border (of length l_i) of the region R_i. The contrast of R_i is :

$$C(R_i) = \begin{cases} 1 - \frac{I_i}{E_i} \text{ if } 0 < I_i < E_i \\ E_i \text{ if } I_i = 0 \\ 0 \text{ otherwise} \end{cases} \qquad (3)$$

The global contrast is :

$$C_{zeboudj} = \frac{1}{A} \sum_i A_i C(R_i). \qquad (4)$$

- Rosenberger's criterion (Rosenberger) : [9]
  The originality of this criterion lies in its adaptive computation according to the type of region (uniform or textured). In the textured case, the dispersion of some textured parameters is used and in the uniform case, gray-levels statistics are computed.
  The contrast function is calculated as follow :

$$C_{Rosenberger} = \frac{\overline{D}(I) + 1 - \underline{D}(I)}{2} \qquad (5)$$

where $\overline{D}(I)$ corresponds to the global intra-class disparity that quantify the homogeneity of each region of image I, and $\underline{D}(I)$ corresponds to the global inter-class disparity that quantify the global disparity of each region of image I.

4. APPLICATION TO PARAMETERS FITTING

We show in this section an application of the previous evaluation criteria for the fitting of parameters of a segmentation method. In general, one has to set the level of precision of the segmentation method by adjusting the parameters of the method to obtain the number of classes or regions in the segmentation result. The goal of this work is to be able to determine automatically the best parameters of a segmentation method to reach a given level of precision. In order to quantify the level of precision of a segmentation result, we choose to consider the number of regions in the segmentation result.

We use for this application the EDISON algorithm [4] and we try to determine the best values of two parameters (spatial and range bandwidth) to segment two aerial images at different levels of precision (see figure 5(a), 6(a)). We segmented these two images with 8 values of the spatial bandwidth (from 3 to 17 with a step of 2) and 11 values of the range bandwidth (from 2 to 31 with a step of 3). So, we obtained 88 segmentation results for each image.

We defined three types of segmentation results by considering the number of detected regions :

- precise segmentation results : segmentation results with a number of regions between 60 and 90,
- intermediate segmentation results : segmentation results with a number of regions between 20 and 40,
- coarse segmentation results : segmentation results with a number of regions between 5 and 10.

We show in table 1 the value of the Zeboudj's contrast computed for the different segmentation results obtained for each value of the two previous parameters on the image presented in figure 5(a) (uniform image).

Range \ Spatial	3	5	7	9	11	13	15	17
2	0,53	0,54	0,54	0,53	0,53	0,53	0,55	0,54
5	0,56	**0,57**	0,56	0,54	0,54	0,53	0,55	0,54
8	0,59	0,55	0,58	0,54	0,52	0,5	0,53	0,54
11	**0,59**	0,52	0,54	0,53	0,51	0,51	0,52	**0.51**
14	0,49	0,49	0,48	0,49	0,49	0,47	0,44	0,44
17	0,01	0,09	0,35	0,47	0,49	0,44	0,44	0,42
20	0,01	0,01	0	0,35	0,35	0,14	0,41	0,39
23	0	0	0	0,30	0,02	0,02	0,07	0,03
26	0	0	0	0	0,03	0,03	0,03	0
29	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0

Table 1: *Evolution of the Zeboudj's contrast for each value of spatial bandwidth (column) and range bandwidth (row) by using the EDISON algorithm on the uniform image.*

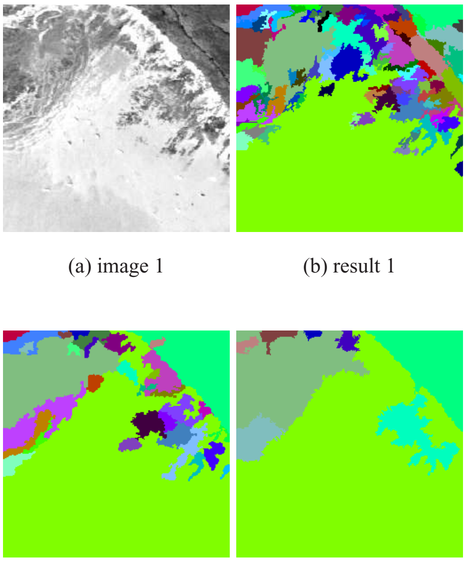
Table 2 gives the number of regions of the different segmentation results. This information allows us to determine the level of precision of each segmentation result.

Range \ Spatial	3	5	7	9	11	13	15	17
2	86	85	82	90	86	82	73	75
5	70	**63**	60	50	46	45	38	37
8	41	41	36	28	35	25	17	24
11	**27**	21	22	20	18	12	13	**10**
14	16	15	13	14	13	10	8	7
17	5	7	8	11	9	5	8	5
20	5	4	4	7	6	3	6	6
23	3	2	2	8	4	5	4	2
26	1	6	1	1	2	2	2	2
29	1	2	1	1	1	1	1	3
31	3	1	1	1	1	1	1	1

Table 2: *Number of regions for each value of spatial bandwidth (column) and range bandwidth (row) by using the EDISON algorithm on the uniform image.*

In order to facilitate the understanding of these tables, we put into obviousness each type of segmentation result with a different gray level (black : precise result, dark gray : intermediate result and light gray : coarse result). When the value of the range bandwidth parameter becomes important, the Zeboudj's contrast has a value equal to 0, meaning that the segmentation result is very bad. That makes sense because in these cases, there are only one or two detected regions.

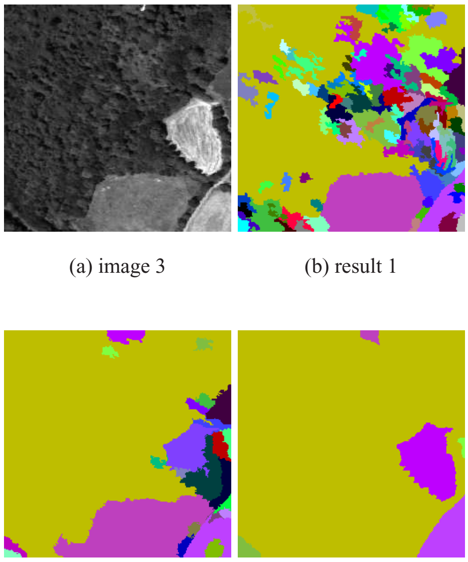
For each type of segmentation result, we choose the parameters maximizing the evaluation criterion adapted for the type of image. For uniform images, we choose the Zeboudj's contrast and for textured images, the Rosenberger's criterion. We present in figures 5, 6 segmentation results obtained by using the parameters maximizing the criterion for each original image and by using the coloring method described in section 2. These segmentation results are visually coherent for each level of precision.



(a) image 1 (b) result 1

(c) result 2 (d) result 3

Figure 5: **Optimal segmentation results of an uniform image :** (a) uniform image - (b) precise segmentation - (c) intermediate segmentation - (d) coarse segmentation.



(a) image 3 (b) result 1

(c) result 2 (d) result 3

Figure 6: **Optimal segmentation results of a textured image :** (a) textured image - (b) precise segmentation - (c) intermediate segmentation - (d) coarse segmentation.

5. CONCLUSION AND PERSPECTIVES

We presented in this paper different techniques that allow the comparison of different segmentation results of an image. We proposed a new visualization method of a segmentation result by a color representation and we showed an application of two evaluation criteria : the Zeboudj's contrast and the Rosenberger's criterion for the determination of the best fitted parameters of a segmentation method to obtain the desired level of precision. Perspectives of this work concern the combination of these two evaluation criteria to improve the reliability of evaluation.

REFERENCES

[1] M.K. Bashar, T. Matsumoto, and N. Ohnishi. Wavelet transform-based locally orderless images for texture segmentation. *Pattern Recongition Letters*, 24(15):2633–2650, November 2003.

[2] B. Bhanu and S. Fonder. Learning based interactive image segmentation. In *IEEE International Conference on Pattern Recognition*, pages Vol I: 299–302, 2000.

[3] S. Chabrier, B. Emile, H. Laurent, C. Rosenberger, and P. Marché. Unsupervised evaluation of image segmentation : Application to multi-spectral images. *IEEE International Conference on Pattern Recognition*, 3:576–579, 2004.

[4] D. Comanicu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

[5] C. Kermad, B. Vozel, and K. Chehdi. Hyperspectral image analysis and dimensionality : a scalar scheme through multi-thresholding technique. *Proc of the Eos/Spie Symposium on Remote sensing*, 31(4170), 2000.

[6] J. MacQueen. Classification and analysis of multivariate observations. *5th Berkeley Symp. Math. Statist. Probability*, pages 281–297, 1967.

[7] T. Ojala and M. Pietikinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32:477–486, 1999.

[8] N. Paragios, Mellina-Gottardo, and Ramesh. Gradient vector flow fast geodesic active contours. In *IEEE International Conference in Computer Vision*, volume 1, pages 67–73, 2001.

[9] C. Rosenberger and K. Chehdi. Genetic fusion : Application to multi-components image segmentation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2219–2222, Istanbul 2000.

[10] R. Zeboudj. *Filtrage, seuillage automatique, contraste et contours : du pré-traitement à l'analyse d'images*. PhD thesis, Saint Etienne University, 1988.

[11] Y.J. Zhang. A survey on evaluation methods for image segmentation. In *Computer Vision and Pattern Recognition*, volume 29, pages 1335–1346, 1996.