Unsupervised Color Image Segmentation Based on Local Fractal Dimension

Karin S. Komati, Evandro O. T. Salles, Mario Sarcinelli-Filho
Graduate Program on Electrical Engineering
Federal University of Espirito Santo
Vitória, Brazil
{kkomati,evandro, mario.sarcinelli}@ele.ufes.br

Abstract— This paper proposes an improved version for the JSEG color image segmentation algorithm, combining the classical JSEG algorithm with a local fractal operator that measures the fractal dimension of each pixel, thus improving the boundary detection. Furthermore, the sensitivity of color variation is enhanced when working with the original color value, instead of quantized color information. Experiments with natural color images of the "*The Berkeley Segmentation Dataset and Benchmark*" (BSDS) are presented, which show improved results, qualitatively and quantitatively, in comparison with the classical JSEG, the Fractal-only and the Fractal-JSEG methods.

Keywords- color image segmentation; J value segmentation (JSEG); local fractal dimension, differential box-counting (DBC).

I. INTRODUCTION

Image segmentation is one of the most important tasks in computer vision. Image segmentation is a process of dividing an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous.

Several methods have been proposed in the literature for color image segmentation. One of the most popular is the one proposed by Deng and Manjunath [1], the JSEG color image segmentation algorithm. It is a very powerful method to test the homogeneity of a given color-texture pattern. However, in some cases it does not perform a high-quality segmentation when compared to the human ground truth provided in BSDS.

In previous works [2], [3], we proposed two different improved versions for the JSEG method: Fractal-JSEG and Fractal-only. We achieve segmentation results that match the human perception better than the segmentation results of the original JSEG by adopting a better way to distinguish inter-regions and intra-regions areas. Both, Fractal-JSEG and Fractal-only, embed the local fractal dimension in a JSEG algorithm, enhancing the detection of boundary regions. However, the Fractal-JSEG method considers the *J*-image, while the Fractal-only algorithm does not.

Despite having improved the JSEG results in [2], [3], some issues remain open. One of such issues is the case of images with background and foreground objects with similar colors and complex textures. Fractal-only approach fails with this kind of images and Fractal-JSEG presents results similar to those of the JSEG.

In this paper, we propose a new approach to solve such problem, enhancing the sensitivity of color variation working with the original color value, instead of a class of this color as JSEG. We continue combining the local fractal dimension in the JSEG algorithm, enhancing the detection of boundary regions, and, as a consequence, the image segmentation results. The new method thus generated is hereinafter called *I-Frac*.

II. PROPOSED METHOD

This section reviews the JSEG method and the local fractal operator used to estimate the fractal dimension (FD) of each pixel. In the following, it presents the proposed architecture mixing the original JSEG algorithm and the FD.

A. JSEG method

The essence of the JSEG method [1] is to separate the segmentation process into two independently processed stages, which are color quantization and spatial segmentation.

In the color quantization process, the colors in the image are reduced through peer group filtering (PGF) and vector quantization. PGF is a nonlinear algorithm for image smoothing and impulsive noise removal. The result of color quantization is a class-map which associates a color class label to each pixel belonging to the class.

In the spatial segmentation process, the criterion to measure the distribution of color classes, *J* measure, is calculated. Essentially, it measures the distances between different classes, divided by the distances between the members within each class, an idea similar to the Fisher's multi-class linear discriminant.

The *J* value can be calculated by using a local area of the class-map, and can indicate if that area is in an interior region or in a boundary region. Thus, a *J*-image, whose pixel values correspond to the J values calculated over local windows centered at the pixels, is built. Multi-scale J-images are calculated changing the local window size.

In the J-image, the higher the local J value is, the more likely the pixel is nearby a boundary region. The J-image is like a 3-D terrain map containing valleys and mountains that actually represent the center regions and boundary regions. The characteristics of J-image allow using a region growing method to segment the image.

As can be seen in Fig. 1, the algorithm starts with a coarse initial scale, and repeats the same processing with the next scale (a smaller window) until the minimum specified scale is reached. Finally, to overcome the oversegmentation problem, regions are merged based on their color similarities, a Euclidean distance measure is applied directly [1].

B. *Local Fractal Dimension*

There are several approaches to estimate the FD in an image. In this work we will use the local differential box-counting method proposed by Vuduc [4] based on the work of Sarkar and Chaudhuri [5].

The differential box-counting method (DBC) is counts the minimum number of boxes of different sizes, which can entirely cover the whole surface, instead of directly measuring an image surface

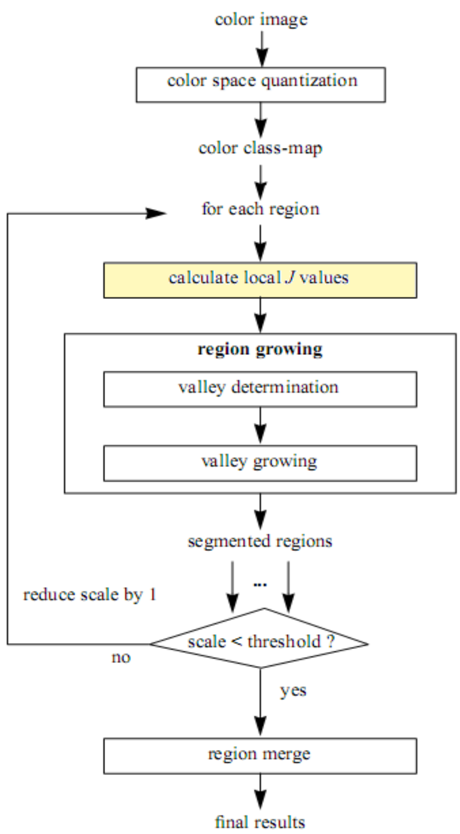


Figure 1. Flow chart of the steps of JSEG (figure modified from [1]).

For a given scale, an image is viewed as a 3-D surface, where it is partitioned into grids. On each grid, there is a column of boxes, where its height is given by the difference between the maximum and minimum gray levels that fall in the grid.

The partition and estimation are performed for different scales, and the multifractal dimension of order q can be estimated for each pixel. The box number counted is an approximation of the optimal one, but very simple and computationally efficient.

Vuduc follows the concept that visual perception of textures can be addressed by analyzing the statistical behavior of the image in a window of limited dimension [6]. Thus, he proposed to measure the fractal dimension (FD) of a single pixel, considering a small window surrounding it.

C. *I-Frac*

In previous works [2], [3], we implemented and tested two approaches: Fractal-only and Fractal-JSEG. Essentially, we change the heart of JSEG method, replacing the J-image calculation by the DBC calculation as shown on Fig. 2. In other words, the new 3-D terrain map is the result of DBC method on original image instead of the process "calculate the local J values" (marked as a filled rectangle with yellow color in the flow chart of Fig. 1).

The "DBC approach". task in Fig. 2 refers to the local fractal dimension. The Fractal images are also a 3D terrain maps, where each pixel represents the FD of the local window. Each FD is converted to be higher in boundary regions and to have the same limits applied to a J-image. That is because the FD in the border regions of a texture is always lower than the FD of the texture as a whole.

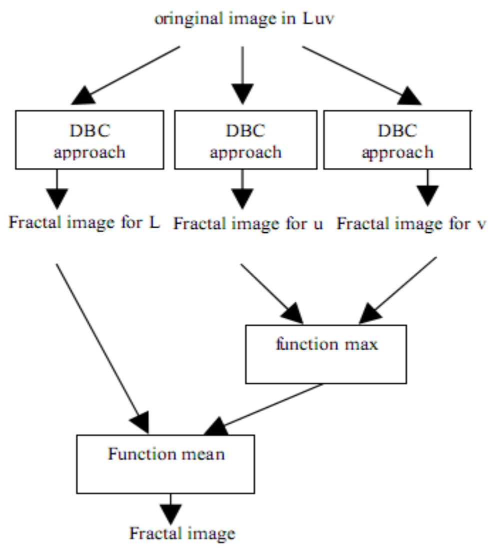


Figure 2. Flow chart of the steps of *I-Frac* image.

Next, we present some differences between the *I-Frac* and the older approaches, JSEG, Fractal-only and Fractal-JSEG:

1) Window Size: The size of local window has been reduced to half of its original size as listed in Table I. The size of local window determines the scale and the smallest scale is denoted scale 1. From scale 1, the window size is doubled each time to obtain the next increased scale. In JSEG method, for computational reasons described in [1], successive windows are downsampled. As downsampling reduces data size, not all pixels in the local window are considered in JSEG, loosing information. In *I-Frac*, there is no longer computational problem as in JSEG, so, there is no downsampling rate, thus improving the accuracy of FD computation.

TABLE I. WINDOW SIZES AT DIFFERENT SCALES

scale	window (pixels)	downsampling (1 / pixels)
1	5 x 5	1 / (1 x 1)
2	9 x 9	1 / (1 x 1)
3	17 x 17	1 / (1 x 1)
4	33 x 33	1 / (1 x 1)

2) Original Image: I-Frac uses as input the original image instead of the output of color quantization to calculate the measurement of homogeneity.

3) Combining Fractal Images: In Fractal-only and Fractal-JSEG, all three 3D terrain maps, from each component of CIELUV color space (Luv), are combined in the *max* function. Preliminary experiments showed that the color components u and v have similar information, so, they are mixed by *max* function. Then, the arithmetic mean of the mixed color map and luminance (L) map is computed.

III. EXPERIMENTAL RESULTS AND DISCUSSION

We tested our *I-Frac* method with natural colored images provided by The Berkeley Segmentation Dataset and Benchmark [7], where human segmented images provide ground truth boundaries. For the presented results, we use all of one hundred images of the test dataset.

Our experiments do not include any parameter-tuning for individual images: the color quantization threshold and the number of scales are chosen automatically as in the original JSEG algorithm. The threshold region merging is the default value (0.4).

The original JSEG algorithm tends to oversegment images, splitting objects into several small regions. This result does not match human perception. Previously in [2], all the segmented images using the Fractal-JSEG algorithm exhibited less segments when compared with the same image segmented using the original JSEG algorithm. The results of *I-Frac* are similar to Fractal-JSEG. However, as the *J*-images are no longer necessary, the algorithm is simplified and the result is that the *I-Frac* method has a lower computational cost when compared to the Fractal-JSEG method.

One problem pointed in [2] is related to images where all elements have almost the same color. For example, the worst case is the image of a snake in the desert. Fractal-only approach was not able to segment any part of the image. Regarding the other methods, Fig. 3 (a) shows the original image, (b) shows the human ground truth, (c) shows the segmentation result of the original JSEG algorithm, (d) shows the segmentation result of the Fractal-JSEG algorithm, (e) and (f) show the segmentation result of I-Frac method with and without the original image included.

The JSEG and Fractal-JSEG present exactly the same result. These segmented regions, however, comprise more shadows than the snake. The segment result of I-Frac is closer to the human ground truth, as it is seen when comparing the parts (b) and (f) of Fig 3.

The quantized snake image presents only two color class maps. For sure, the color quantization method causes loosing color information. Thus, to achieve this result, the use of original image instead of the output of a color quantization step was essential.

The size of local window determines the size of image regions that can be detected. Windows of small size are useful in localizing the intensity/color edges. The reduction in the local window size in *I-Frac* improves the sensitivity to edges without increasing the number of segmented regions.

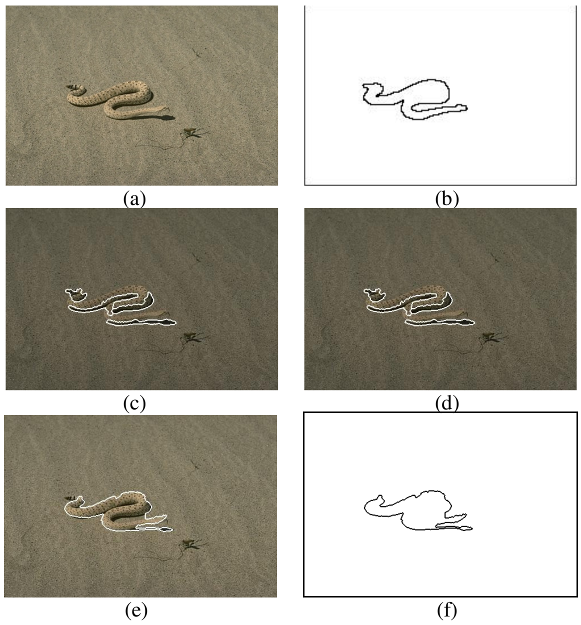
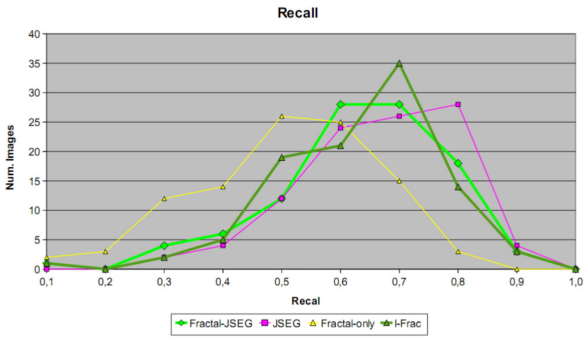


Figure 3. (a) Original Image (b) Human benchmark (c) Result of JSEG method (d) Result of Fractal-JSEG method (e) Result of I-Frac method (f) Result of I-Frac without the original image.

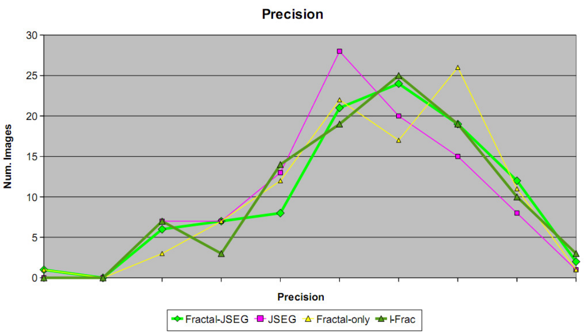
In order to evaluate quantitatively the models against each other, we compute three metrics: precision, recall and F-measure [8]. Our segmentation results are hard edge maps, i.e. binary images, where "1" marks the segment boundary pixels. For this reason, we do not work at many levels of threshold, so we do not use the traditional precision-recall curve. The graphs presented in Fig. 4 are histograms. Each point in the graph represents the number of images with (a) recall (b) precision (c) F-measure falling into the interval that the point is located on. So, we can read in Fig. 4a that I-Frac presented 35 images (from the total 100 images) with recall=0.7.

As we can see in the graphics of Fig. 4, JSEG presents a better recall but it has the worst precision results. The Fractal-only method shows better performance in precision but fails in recall. The Fractal-JSEG and *I-Frac* present results between JSEG and Fractal-only. Both of them present better results than Fractal-only in recall, but worse than JSEG. *I-Frac* is better than Fractal-JSEG because the global maximum is largest than the global maximum of Fractal-JSEG for a high value of recall=0.7. Same analysis can be applied to precision graphics, where *I-Frac* is slightly better than Fractal-JSEG.

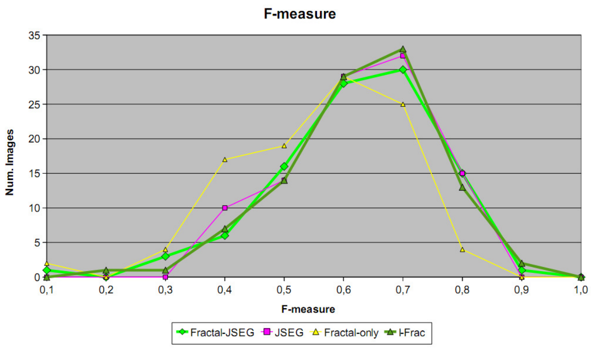
Considering that we reached a trade-off between precision and recall, i. e. improved one in detriment of the other; the graphs forms obtained for F-measure of all methods are very similar (see Fig. 4c). The *I-Frac* F-measure is better than all, and, again, with a little difference between Fractal-JSEG. We want to emphasize that *I-Frac* method increases the F-measure a bit, considering all the test dataset of BSDS, the problem related to images where the elements have almost the same color was resolved.



(a)



(b)



(c)

Figure 4. Comparison of Fractal-JSEG x JSEG x Fractal-only x *I-Frac* (a) Recall (b) Precision (c) F-measure.

IV. CONCLUSIONS

We propose an improved version for the classical JSEG algorithm. Our technique integrates the classical JSEG algorithm and the local fractal operator that measures the FD of each pixel, thus improving boundary detection. Moreover, we enhance the sensitivity of color variation working with the original image instead of quantized output. In that way, the technique provides a solution to the poor segmentation results of images with background and foreground objects having similar colors and complex textures.

The conclusion is that the *I-Frac* approach improves the sensitivity to boundary regions, thus providing segmentation results that match the human perception better than the segmentation results associated to the original JSEG algorithm, the Fractal-only method and the Fractal-JSEG method.

ACKNOWLEDGMENT

The authors would like to thank CAPES (Brazil) for financial support and PPGEE (UFES) for research development support.

REFERENCES

[1] Y. Deng and B. S. Manjunath, “Unsupervised segmentation of color-texture regions in images and video”, *IEEE Trans. Pattern Anal. Mach. Intell.,* vol 23, no 8, pp. 800–810, August 2001.

[2] K. S. Komati, E. O. T. Salles, and M. Sarcinelli-Filho, “Fractal-JSEG: JSEG using a homogeneity measurement based on local fractal descriptor,” in *Proc. of the XXIInd Brazilian Symposium on Computer Graphics and Image Processing. SIBGRAPI 2009*. October 2009.

[3] K. S. Komati, E. O. T. Salles, and M. Sarcinelli-Filho, “Unsupervised color image segmentation based on local fractal descriptor an J-images,” in *Proc. of the IEEE ICIT 2010 International Conference on Industrial Technology*, vol. 1, March 2010, in press.

[4] R. Vuduc, “Image segmentation using fractal dimension,” 1997. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.132.6547

[5] N. Sarkar and B. Chaudhuri, “An efficient differential box-counting approach to compute fractal dimension of image,” IEEE Transactions on Systems Man and Cybernetics, vol. 24, no. 1, pp. 115–120, 1994.

[6] R. Manduchi and J. Portilla, “Independent component analysis of textures,” in *Proc. of the 7th IEEE International Conference on Computer Vision*, vol. 2, 1999, pp. 1054–1060.

[7] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proc. 8th IEEE Int'l Conf. Computer Vision*, vol. 2, July 2001, pp. 416–423.

[8] D. R. Martin, C. C. Fowlkes, and J. Malik, “Learning to detect natural image boundaries using local brightness, color, and texture cues,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 5, pp. 530–549, 2004.