Fractal-JSEG: JSEG Using an Homogeneity Measurement Based on Local Fractal Descriptor

Karin S. Komati, Evandro O. T. Salles, Mario Sarcinelli Filho
Graduate Program on Electrical Engineering
Federal University of Espirito Santo
Vitoria, 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 and a local fractal operator that measures the fractal dimension of each pixel, thus improving the boundary detection in the J-map. Experiments with natural color images of the Berkeley Segmentation Dataset and Benchmark are presented, which show improved results in comparison with the classical JSEG algorithm.**

***Keywords*-color image segmentation, J value segmentation (JSEG), edge detection, fractal dimension, local fractal operator, differential box-counting (DBC).**

I. Introduction

Image segmentation is one of the most important tasks in computer vision. Its objective is to separate one image into disjoint homogeneous regions compatible with human perception. As a matter of fact, high level procedures, like object recognition, strongly rely on the quality of image segmentation. Image segmentation has been investigated for the last thirty years, still remaining as a hard problem.

Several methods have been proposed in the literature of image segmentation. One of the most popular is the one proposed by Deng and Manjunath [1], [2], the JSEG algorithm. It is a very powerful method to test the homogeneity of a given color-texture pattern, and is quite efficient in computation terms. However, in some cases it does not perform a high-quality segmentation.

Several improvements on the JSEG algorithm have been proposed. Chang et al. [3] proposed an improved contrast method (IC-JSEG) that considers the color information of the original pixels instead of its class-map. Experiments with natural images show that such method is robust and produces better results than the JSEG one. Yu and colleagues [4] worked on a similar idea to improve results. Wang et al. [5] presented an extension to the JSEG algorithm that integrates directional operators to improve the measurement of the texture structure.

We believe that JSEG can mitigate segmentation problems by adopting a better way to distinguish inter-regions and intra-regions. Edge detection operators are not suitable to describe the textural homogeneity, while fractals are. Indeed, natural images are well-represented by statistical fractals [6]. The statistical self similarity of the fractals refers to the fact that the statistical measurement of a signal is invariant to scale transformation. In this sense, a fractal surface is one that can be precisely approximated by a simple fractal function, over a range of scales. They also incorporate high-order statistical information in the texture representation, through the spatial frequency indirectly inserted in the fractal dimension. This is in agreement with Gagalowicz (according to a citation by Pratt [7]), who has shown that high-order statistics is necessary to classify textures. Thus, fractals represent very well natural images, since they do not change their characteristics over different scales and also do incorporate high-order statistics.

Visual perception of textures can be addressed by analyzing the statistical behavior of the image in a window of limited dimension [8]. This can be perceived through using statistical filters that maps the image values in a certain neighborhood in a subspace of perceptibility, thus reducing the size of the representation while preserving the structural information. The JSEG algorithm already implements the idea of a local window to compute its own homogeneity criterion. In this paper, we follow this concept, computing a local fractal operator that can measure the fractal dimension (FD) of a single pixel, considering a small window surrounding it [9].

The fractal dimension in the border regions of a texture is always lower than the fractal dimension of the texture as a whole. Thus, using the FD it is possible to determine the border lines separating regions of different textures, as it is shown in [10], [11], [12]. Côco, Salles and Sarcinelli-Filho [10], [11] adopted a fractal dimension to describe texture in a TICA model applied to image segmentation with good results. Conci and Nunes [13] also used this method as the base for an efficient computational algorithm.

In this paper, we propose a new idea by embedding the local fractal dimension in a JSEG algorithm, enhancing the detection of boundary regions, and, as a consequence, the image segmentation results. Moreover, we enhance the sensitivity of color variation working with the original value of color, instead of a class of this color. The new method thus generated is hereinafter called Fractal-JSEG.

The rest of this paper is organized as follows: Section II reviews the JSEG method and makes clear why to use the real color instead of class-map, while Section III describes the local fractal operator used to estimate the fractal dimension of each pixel in each scale. In the sequel, Section IV presents the proposed architecture mixing the original JSEG algorithm and the fractal dimension. Following, Section V presents experimental results that demonstrate that our approach improves the original method. Finally, Section VI highlights our conclusions and some future works.

II. The JSEG Method

The essence of the JSEG method is to separate the segmentation process into two independently processed stages, which are color quantization and spatial segmentation, according to the schematic shown in Fig. 1. These stages are described as:

1. Color Quantization
Colors in the image are reduced through peer group filtering (PGF) [14] 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.

2. Spatial Segmentation
J measure is the criterion to measure the distribution of color classes and is defined as follows: let Z be the set of all N data points in the class map, $z = (x, y)$, $z \in Z$, and m be the mean. Suppose that Z is classified into C classes, $Z_i, i = 1, ..., C$. To N_i data points of class Z_i, we can write

$$m_i = \frac{1}{N_i} \sum_{z \in Z_i} z. \tag{1}$$

Let us denote S_T as the total variance of data points in Z, which is given by

$$S_T = \sum_{z \in Z} \|z - m\|^2, \tag{2}$$

and S_W as the total variance of the points belonging to the same class, which is defined as

$$S_W = \sum_{i=1}^{C} S_i = \sum_{i=1}^{C} \sum_{z \in Z} \|z - m_i\|^2. \tag{3}$$

Then, the J value is defined as

$$J = S_B / S_W = (S_T - S_W)/S_W. \tag{4}$$

Essentially, (4) measures the distances between different classes, S_B, divided by the distances between the members within each class S_W, an idea similar to the Fisher's multi-class linear discriminant [15]. A higher value of J indicates that the classes are more separated one from each other and the members within each class are closer one to each other, and vice versa.

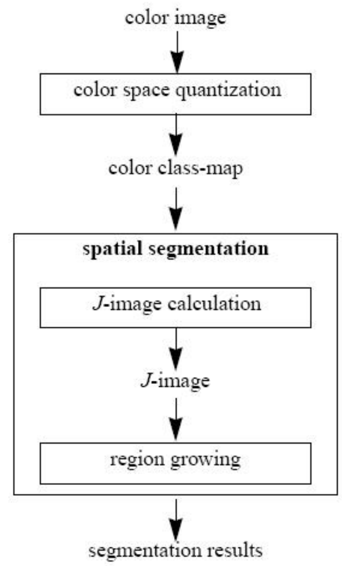


Figure 1. Schematic of the JSEG algorithm [2]

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, the size of the local window determining the size of the image regions that can be detected. 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 to use a region growing method to segment the image. 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.

The results produced by the JSEG method are mainly based on the class-map produced in the first step. The class-map is formed by numbers from 1 to the number of classes. For example, if quantization reduces to 10 colors, then the class-map has numbers between 1 to 10, instead of the average color value of each class. The measure J is defined on the variance of this class-map, which describes the texture information, not considering color information about the pixels. As an example, assume that the class color

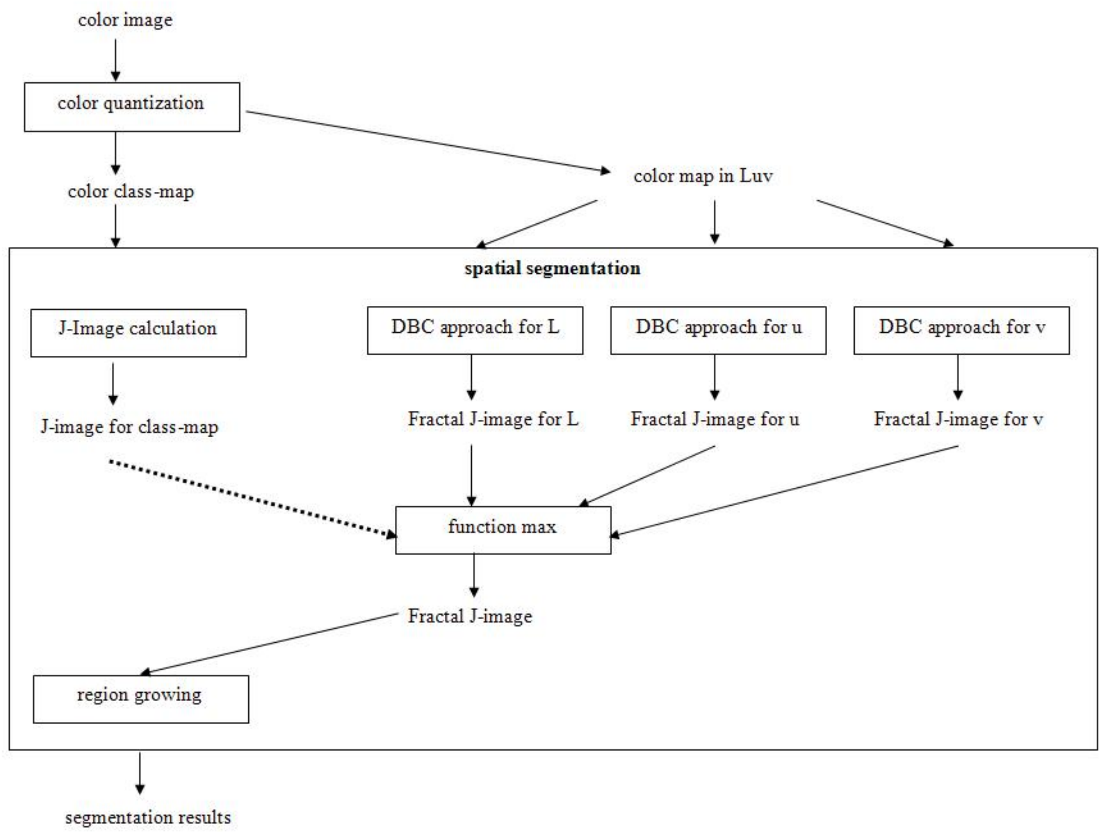


Figure 2. Schematic of the Fractal-JSEG algorithm.

1 is light blue, the class color 2 is dark blue and the class color 3 is red. The variance based on true color will be higher between red and blue than between light and dark blues. However, when using the classical JSEG method the J values of classes color 1 and 2 or 2 and 3 are basically the same. So, to obtain more sensitivity to colors, our approach works with the original image color. The drawback of such method is that it tends to generate oversegmentation [3].

III. THE LOCAL FRACTAL MEASURE

There are several approaches to estimate the FD in an image. In this work we will use the box-counting method proposed by Sarkar and Chaudhuri [16]. Vuduc [9] applies three different local FD operator techniques - threshold, Sarkar's and Beaver's methods - to a series of images. He discovered that the Sarkar method is surprisingly effective on a number of different kinds of images. Vuduc also introduces the notion of a local fractal operator that can "measure" the fractal dimension of a single pixel, which is used in this work.

The method proposed by Chadhouri and Sarkar is based on the differential box-counting (DBC) algorithm. Instead of directly measuring an image surface, the measure is obtained by means of counting the minimum number of boxes of different sizes, which can entirely cover the whole surface. The process is detailed as follows: for a given scale ε, an $M \times M$ image is partitioned into grids of size $\varepsilon \times \varepsilon$. On each grid, there is a column of boxes, where $\varepsilon' = [\varepsilon \times G/M]$ and G is the maximum gray level. The image is viewed as a 3-D surface, where (i, j) denotes the 2-D position and the third coordinate, z, denotes the gray level of the corresponding pixel. Given the maximum and minimum gray levels in the $(i, j)^{th}$ grid that fall in the v^{th} and u^{th} boxes, respectively, the number of boxes, η_ε, needed to cover the image surface on that grid is calculated as

$$\eta_\varepsilon(i, j) = v - u + 1, \tag{5}$$

and the total number of boxes, $\tilde{N}_\varepsilon$, needed to cover the whole surface can then be estimated as

$$\tilde{N}_\varepsilon = \sum_{i,j} \eta_\varepsilon(i, j). \tag{6}$$

In order to describe the distribution of different subfractals, a measure $\mu_\varepsilon(i, j)$ is defined on the grid as

$$\mu_\varepsilon(i, j) = \frac{\eta_\varepsilon(i, j)}{\tilde{N}_\varepsilon}. \tag{7}$$

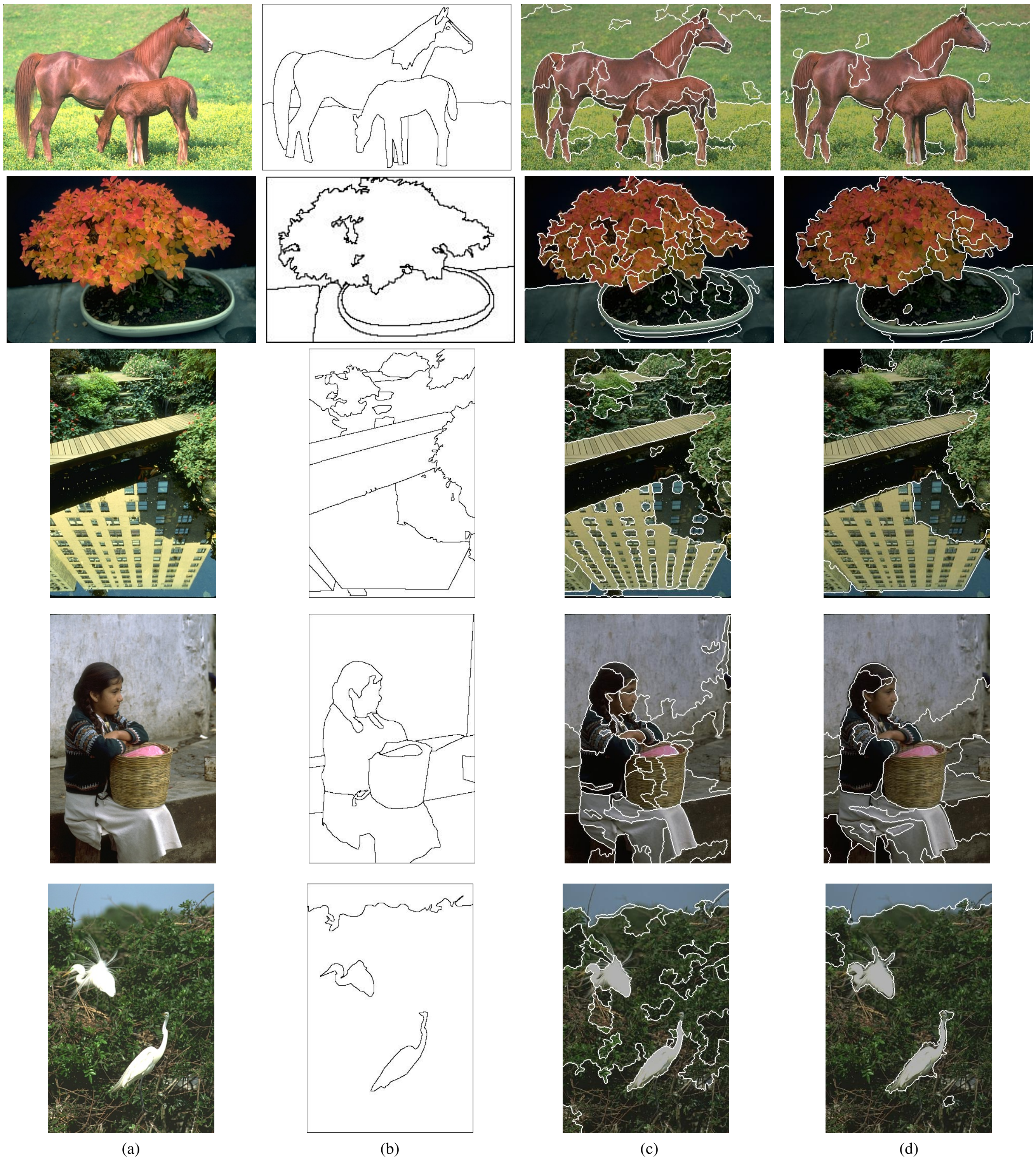


Figure 3. (a) Original image (b) Human benchmark (c) Results of the JSEG method (d) Results of the fractal-dimension-only method.

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

Figure 4. (a) Original image (b) Human benchmark (c) Results of the JSEG method (d) Results of the fractal-dimension-only method.

The partition and estimation are performed for different scales, and the multifractal dimension of order q can be estimated for each pixel, which is given by

$$D_q(i,j) = \frac{1}{1-q} \lim_{\varepsilon \to 0} \frac{\ln\left[\mu_\varepsilon(i,j)^q\right]}{\ln\left(\frac{1}{r}\right)}, \quad (8)$$

where $r = \varepsilon/M$.

The box number counted is an approximation of the optimal one, but very simple and efficient.

IV. The Proposed Method

We propose a new color image segmentation approach based on the conclusion presented in [11], [10] (fractal dimension models texture quite well) and in [3] (color information of original pixels is better than the class-map values).

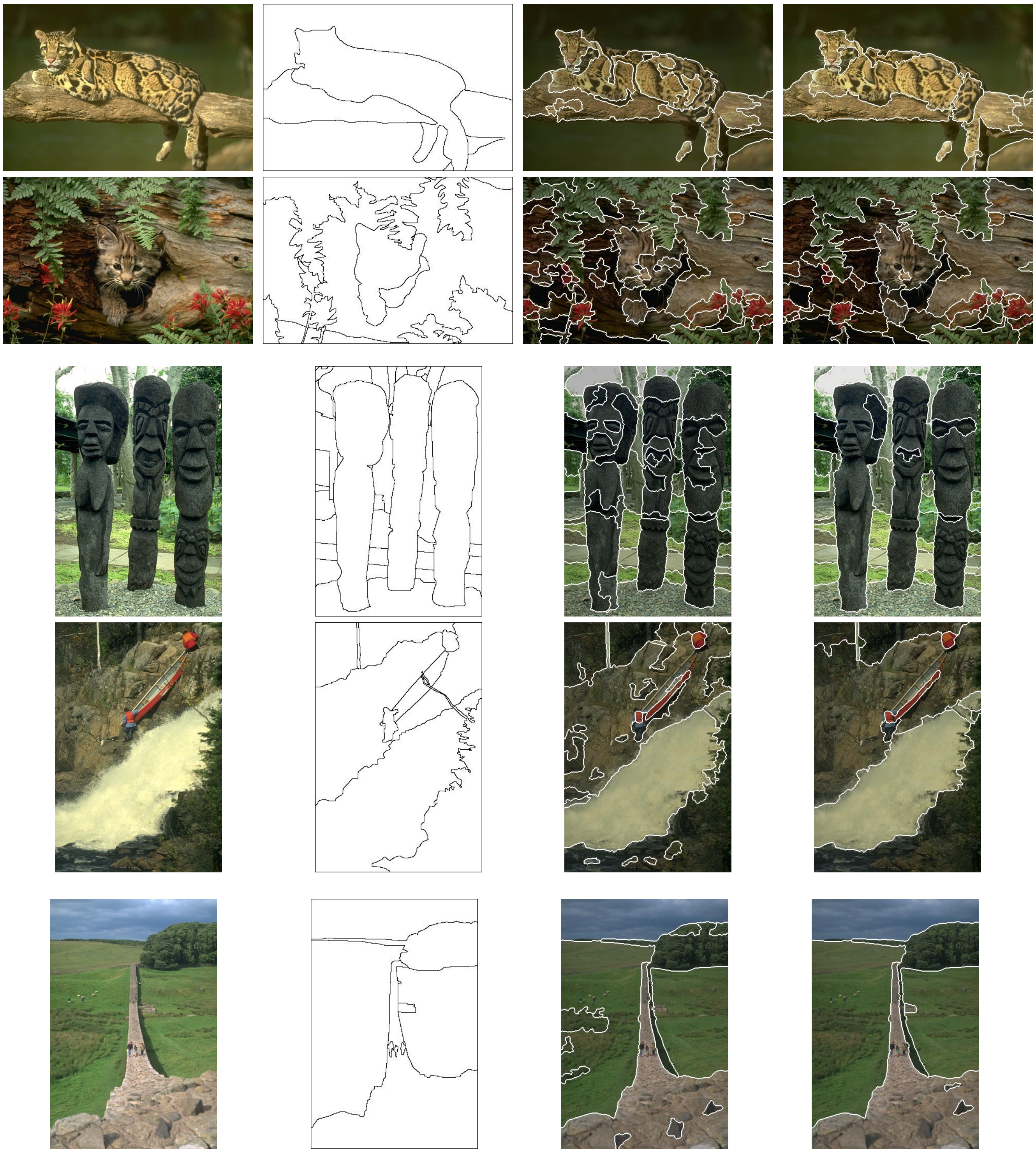
During this research, we implemented and tested two approaches. The first one computes the 3D terrain map using only the local fractal operator on color-map in Luv format. The second one mixes the first attempt and the original J-image. Fig. 2 describes these two software architectures. The first one was designed as the sequence defined by the solid-line arrows, while the sequence correspondent to the second one includes all the arrows. The new method, called Fractal-JSEG method, corresponds to the second one. Notice that the Fractal-JSEG proposal keeps the color quantization and region growing processes of the classical JSEG method.

The "DBC approach" task in Fig. 2 refers to the local fractal measure. The Fractal-JSEG image is also a 3D terrain map, 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. The local window used to compute FD has the same size as the local window used to compute the J-image.

The Fractal-JSEG images compute each Luv component separately. All 3D terrain maps are combined in the max function. This final terrain map is higher than or equal to any one of the three original components in each pixel. If the fractal dimension of one pixel in the component u is higher than the fractal dimensions of the same pixel in the component L, as an example, it means that this pixel is a boundary pixel more perceivable in a color component than in the gray component. Taking the maximum value of each pixel, we obtain more definition of mountains and smaller valleys. The decreased size of the valleys decreases the number of seeds used in the region growing, thus decrease the number of regions.

Since in the JSEG method the Luv color space is adopted, the perceptually uniform Luv color space is used to keep consistence. Moreover, preliminary tests with RGB color space presented worse results in comparison with Luv segmentations.

The segmentation results with only the three Fractal-JSEG images shows good improvements compared to the JSEG method. The best results can be seen in Fig. 3, where image (a) shows the original image, image (b) shows the human



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

Figure 5. (a) Original image (b) Human benchmark (c) Results of the JSEG method (d) Results of the Fractal-JSEG method.

ground truth that best matches our final result, image (c) is the segmentation result of the original JSEG and image (d) is the segmentation result considering only the fractal dimension (the first software architecture represented in Fig. 2). One can verify that this approach generates less regions, compared to the original JSEG. Images with different colored regions, specially foreground and background colors, presented the best results.

However, as one can see in the results of Fig. 4, some results were worse than the results of the original JSEG method. In particular, the image with a snake and the image with a goat illustrate such remark. That approach (fractal dimension only) was not able to segment neither the snake nor the goat. In the first case, even a single region was segmented. This is the worst case, because when the segmentation output is inputted to a higher level process (e. g., object recognition), it is necessary to have at least a single segmented region (oversegmentation, although not desirable, is better than no segmentation at all). We analyzed all segmentations with results worse than original JSEG results, and noticed that the common characteristic is that the colors of the image elements are almost the same. An important remark here is that we have not used any criteria to choose the human benchmark in Fig. 4.

Quattrochi and Goodchild [17] states that most real-world surfaces are not perfect fractals. In addition, Pratt and Faugeras [7] show that the human vision system is also sensitive to the correlation between pairs of textures, besides being sensitive to differences in the mean and in the variance. The J-image concept follows this idea of mean and variance. According to the results of our tests, only the fractal dimension does not model the segmentation problem perfectly, and the J-image is very important to identify different textures with similar colors.

Combining both measures, the local fractal dimension of each component of Luv color space and the homogeneity measure of texture-color, we obtained a better classifier (the Fractal-JSEG one), which is more stable and more generic. The results of our tests with such method are presented in the sequel.

V. Experimental Results

We tested our Fractal-JSEG method with natural colored images provided by Berkeley Segmentation Database [18], where human segmented images provide ground truth boundaries.

To test the generalization, 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, and the region merging threshold is the default value (0.4). Other works set the quantization parameter to specific values, (it is set to 200 in [4] and to 150 in [5]). Finally, the order q is set to 2 in the FD computation.

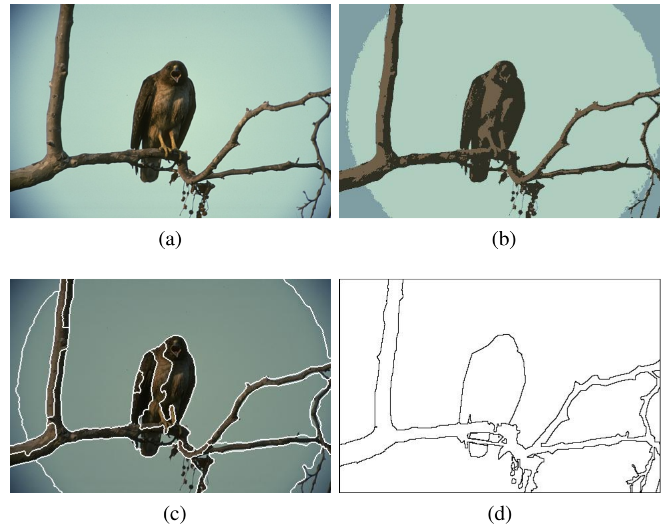


Figure 6. (a) original image (b) quantized image (c) segmented image (d) human segmentation.

The results are compared with the segmentation results of the original JSEG algorithm. Fig. 5 shows the best segmentation results obtained by the Fractal-JSEG method when compared to human segmentation. There (a) shows the original image, (b) shows the human ground truth that best matches our result, (c) shows the segmentation result of the original JSEG algorithm and (d) shows the segmentation result of the Fractal-JSEG algorithm. One can see that the results obtained with the Fractal-JSEG algorithm are closer to the human ground truth than the results associated to the original JSEG algorithm.

The original JSEG algorithm tends to oversegment images, splitting objects into several small regions. This results does not match human perception. For example, in the image with three statues, one human can perceive each statue as a whole and do not segment it. The same concept can be applied to the feline on a tree: a human being understands the animal contour and not several textures in the animal. As a matter of fact, all the images segmented using the Fractal-JSEG algorithm exhibited less segments when compared with the same image segmented using the original JSEG algorithm. This is a clue that oversegmentation is not a problem for the proposed Fractal-JSEG method.

For sure, any color quantization method causes loosing color information. Nevertheless, in some images such losses become a problem when segmenting the image. These difficulties are still bigger when color varies smoothly in certain image regions. As an example, see Fig. 6, where there is a color circle in the quantized image, inducing the idea of a boundary, while the benchmark, i.e., the human segmentation, does not perceive this as a different region.

VI. Conclusion and Future Work

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 fractal dimension of each pixel, thus improving boundary detection in the J-map.

Even when the human beings who generated the ground truth for the database used agree in terms of edges, they often disagree in terms of precisely which pixels in an image correspond to such edges. To account for this, in the future we intend to work on soft edge maps, which will be evaluated using precision-recall graphics. The Fractal-JSEG results presented here are the so called hard edge map because every pixel is determined to belong or not a boundary.

Segmentation is a hard problem and the comparison with other methods is even harder. Using precision-recall curve, it will be possible to translate the performance of an algorithm into a single number: the maximum F-measure. F-measure will provide a quantitative comparison with other methods.

Some improvement in the color quantization step is also necessary to resolve the drawback showed on Fig. 6, and will be addressed in the sequence of this research.

Another future work is to elaborate a more intelligent heuristic to combine the values of the two different measurements (FD and J value), instead of using the max function. As a more intelligent predicate, we could use, for instance, a weighted sum to combine the fractal dimension and the J value.

In spite of such possible improvements, the conclusion is that the proposal here presented 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.

Acknowledgments

The authors would like to thank CAPES (Brazil) for the scholarship given to Mrs. Komati and Dr. Klaus Fabian Côco for the initial DBC implementation.

References

[1] Y. Deng, B. S. Manjunath, and H. Shin, "Color image segmentation," in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'99*, vol. 2. Los Alamitos, CA, USA: IEEE Computer Society, Jun 1999, pp. 446–51.

[2] 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.

[3] Y. C. Chang, J. K. Archibald, Y. G. Wang, and D. J. Lee, "Texture-based color image segmentation using local contrast information," *International Journal of Information Technology and Intelligent Computing*, vol. 2, no. 4, November 2007.

[4] S. Y. Yu, Y. Zhang, Y. G. Wang, and J. Yang, "Unsupervised color-texture image segmentation," *Journal of Shanghai Jiaotong University (Science)*, vol. 13, no. 1, pp. 71–75, February 2008.

[5] Y.-G. Wang, J. Yang, and Y.-C. Chang, "Color-texture image segmentation by integrating directional operators into jseg method," *Pattern Recogn. Lett.*, vol. 27, no. 16, pp. 1983–1990, 2006.

[6] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 661–674, nov 1984.

[7] W. Pratt, *Digital Image Processing (Second Edition)*. Wiley, 1991.

[8] 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.

[9] R. Vuduc, "Image segmentation using fractal dimension," 1997. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.6547

[10] K. F. Côco, E. O. T. Salles, and M. Sarcinelli-Filho, "Topographic independent component analysis based on fractal theory and morphology applied to texture segmentation," *Signal Process.*, vol. 87, no. 8, pp. 1966–1977, 2007.

[11] K. F. Côco, E. O. T. Salles, and M. Sarcinelli-Filho, "Topographic independent component analysis based on fractal and morphology applied to texture segmentation," *Lecture Notes in Computer Science*, vol. 5441, pp. 491–498, 2009.

[12] B. B. Chaudhuri and N. Sarkar, "Texture segmentation using fractal dimension," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 1, pp. 72–77, 1995.

[13] A. Conci and E. O. Nunes, "Multi-bands image analysis using local fractal dimension," in *Proc. of the XIV Brazilian Symposium on Computer Graphics and Image Processing*, vol. 0. Los Alamitos, CA, USA: IEEE Computer Society, 2001, pp. 91–98.

[14] Y. Deng, C. Kenney, M. Moore, and B. S. Manjunath, "Peer group filtering and perceptual color image quantization," in *Proc. of the 1999 IEEE International Symposium on Circuits and Systems. ISCAS '99.*, vol. 4, 1999, pp. 21–24.

[15] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. New-York, USA: John Wiley and Sons, Inc, 2001.

[16] 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.

[17] D. A. Quattrochi and M. F. Goodchild, *Scale in remote sensing and GIS*. CRC Press, 1997.

[18] 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.