

# Simplified Quadtree Image Segmentation for Image Annotation

Gerardo Rodmar Conde Márquez  
Benemérita Universidad Autónoma de Puebla  
14 sur y Av. San Claudio, San Manuel, 72570, Puebla, Mexico  
condemarquezg@acm.org

Hugo Jair Escalante,  
Universidad Autónoma de Nuevo León,  
Graduate Program in Systems Engineering  
San Nicolás de los Garza, NL 66450, Mexico  
hugo.jair@gmail.com,

Luis Enrique Sucar  
Instituto Nacional de Astrofísica Óptica y Electrónica  
Luis Enrique Erro 1, Tonantzintla, 72840, Puebla, Mexico  
esucarg@ccc.inaoep.mx

## Abstract

This paper introduces a *Quadtree* image segmentation technique to be used for image annotation. The proposed method is able to efficiently divide the image in homogeneous segments by merging adjacent regions using border and color information. Our method is highly efficient and provides segmentations of acceptable performance; the segments generated with the proposed technique can be used for automatic image annotation and related tasks (e.g., object recognition). A benefit of our proposal is that it allows us finishing the segmentation at any time by controlling the desired level of the detail for the segmentation; hence, the method is suitable for time restricted scenarios. We compare the performance of an image annotation technique trained on hand labeled images and tested in images segmented with different segmentation algorithms. We found that the best results were obtained when the annotation method was tested in images segmented with the quadtree formulation. Our results give evidence of the efficiency and effectiveness of the proposed method.

## 1 Introduction

Although it is possible to annotate images without a previous segmentation, in general region-level image annotation results in labels of better quality [1, 2]. The latter is due to the fact that region-level information brings additional information (e.g., shape information or spatial context knowledge) that is not present in image-level annotation [1, 2, 3, 4]. Of course, the better the segmentation quality the better is expected to be the performance of region labeling techniques. However, despite the fact that important advances have been reported on image segmentation, most segmentation algorithms tend to fall in one of two extremes. On the one hand, there are sophisticated techniques that tend to give better results, but are too time consuming and thus not appropriate for some applications such as image retrieval [5]. On the other hand, there are methods that are simple and fast (e.g., grid segmentation) although they do not provide a good support for automatic annotation [6]. An attractive alternative consists of methods that offer a tradeoff between efficiency and segmentation quality. This paper presents one of such alternative segmentation techniques that aims at give support to automatic image annotation and content-based image retrieval (CBIR).

---

Luis Enrique Sucar and Hugo Jair Escalante (eds.): AIAR2010: Proceedings of the 1<sup>st</sup> Automatic Image Annotation and Retrieval Workshop 2010. Copyright ©2011 for the individual papers by the papers' authors. Copying permitted only for private and academic purposes. This volume is published and copyrighted by its editors., volume 1, issue: 1, pp. 24-34

The proposed method for image segmentation is a simplified quadtree technique based on the following guidelines: recursively divide the image using a quadtree approach, merge homogeneous and similar quadtrees' regions based on borders and color information, process and discard large image regions as fast as possible, prefer detection of large objects, avoid noise in result segments and provide *any-time* segmentation. Our proposal implements the mentioned guidelines by dividing the segmentation process in five steps: (1) edge detection, (2) border processing, (3) color discretization, (4) quadtree scanning and (5) segmentation enhancement. Although our formulation is intended for automatic image annotation usage, it can be used for other tasks that require a fast segmentation implementation. We present preliminary results on the application of the proposed segmentation algorithm in the task of image annotation. Experimental results show that the proposed technique can be very useful for this task and motivates research in several directions.

The rest of this paper is organized as follows. The next section presents background information on quadtree image segmentation. Section 3 introduces the proposed method. Section 4 describes the experimental settings we adopted for the evaluation of the proposed segmentation method. Section 5 reports experimental results obtained from our formulation. Finally, Section 6 presents some conclusions and outlines future work directions.

## 2 Preliminaries

This section presents background segmentation on quadtree image segmentation that will be helpful for understanding the rest of the paper.

### 2.1 Image segmentation

According to T. Pavlidis the segmentation of an image  $I$  into a set of regions  $S = \{S_1, \dots, S_k\}$  should satisfy the following four main rules [10]:  $S$  is a partition of  $I$  that covers the whole image, no regions intersect, homogeneity predicate is satisfied by each region and union of adjacent regions does not satisfy it. This rules can be expressed as follows:

$$\cup_{i=1}^k S_i = S \quad (1)$$

$$S_i \cap S_j = \phi, \forall i \neq j \quad (2)$$

$$\forall S_i, P(S_i) = true \quad (3)$$

$$P(S_i \cup S_j) = false, \forall i \neq j, S_i adjacent S_j \quad (4)$$

where  $P(X)$  represents the fulfilment of the homogeneous predicate under  $X$ . However, achieving this type of segmentation involves many difficulties. The main issues are due to borders, textures and illumination changes that make hard to distinguish where objects start and end. Moreover, segmenting an image is a highly subjective process. Therefore, it is difficult to make everybody agree on one segmentation or another. Aside these difficulties, most image segmentation algorithms are very time consuming. For these reasons there is a research trend on image annotation that studies annotation methods that avoid the use image segmentation [11, 12]. Nevertheless, previous work on image annotation has revealed that region-level image annotation can result in labels of better quality than those generated with image-level techniques [2, 3, 7, 8]. Therefore, we think it is crucial to improve the performance of image segmentation methods in support of annotation techniques.

## 2.2 Weak segmentation

L. T. Lan and A. Boucher presented a simplified image segmentation process called *Weak Segmentation* [6]. They introduce this method as an easy way of getting over common difficulties of the segmentation process. It avoids generation of perfect segments. Instead, it tries to figure out -in a general way- where and what are the main objects of the image. This approach reduces segmentation costs in effort and time. Figure 1 shows an example of weak segmentation. Based on the weak segmentation formulation, we propose a quadtree based algorithm capable of segmenting images very quickly, yet able to provide segments which can be used for automatic image annotation.



Figure 1: Weak Segmentation example. Left.- Original image. Right.- Segmentation result. Even when the result segments are not perfect we obtain a general idea of which and where are the main regions

## 2.3 Quadtrees

The *Quadtree* structure was first introduced by Hanan Samet in [13]. A *Quadtree* is a data structure concept that refers to a hierarchical collection of maximal blocks that partition a region. All blocks are disjoint and have predefined sizes according to the base quad size. The item to be partitioned is the root quadtree which is recursively partitioned according to predefined criteria. Each step of decomposition produces four new quadtrees of the same size that are hierarchically associated with their parent quadtree. Decomposition finishes whenever there are no more quadtrees to be partitioned or when the quadtrees have reached their minimum size. Quadtree based approaches are commonly used due to their ability to discard very quickly large amounts of information.

## 3 Simplified quadtree image segmentation method

This section describes the proposed technique for image segmentation, which is based on the idea of *Dividing the image following a quadtree structure and merging similar adjacent regions*. The proposed algorithm for simplified quadtree image segmentation requires the specification of three parameters i) minimum object size (*mos*), ii) minimum quad size (*mqs*) and iii) homogeneity threshold (*ht*). These parameters can be adjusted by the user according to their own needs, although below we provide default values for those parameters. The proposed formulation is divided into the following five stages: (1) edge detection, (2) border processing, (3) color discretization, (4) quadtree scanning and (5) segmentation enhancement. Figure 2 shows the implementation design model. The rest of this section provides a detailed description of each step.

### 3.1 Border Detection

The first step consists of detecting borders of objects in the image. Borders are important for image segmentation because they can provide information about object contours. In our proposal, border in-

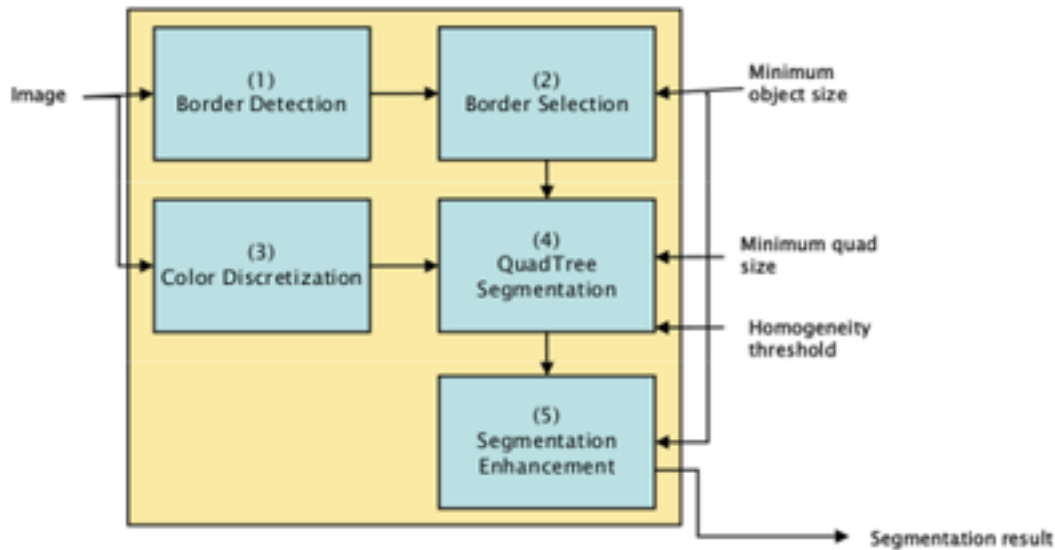


Figure 2: Functional model of our proposal.

formation has a crucial influence over the final segmentation results. There exist many border detection algorithms that provide very good results but they are complex and time consuming. Therefore, we decided to use the Sobel operator which is a simple and straight forward method to detect image borders [9]. A threshold is applied to the results obtained with the Sobel operator in order to keep only the highest values.

### 3.2 Border Selection

Border selection consists of building an image that contains only the most relevant borders obtained from step 1. Border relevance is measured by its continuity and length. To be able to evaluate these measures, we first apply the connected components algorithm to the Sobel result and then we calculate the area of each component. We keep all the components whose area suggest the they belong to an object that reaches the *mos* parameter entered to the algorithm. We call this image *Constraints Image*.

### 3.3 Color Discretization

Color discretization creates a 6-bit color copy of the image. 2 bits per each RGB component. We choose this small number to speed ups histograms calculations as normalizing and comparing.

### 3.4 Quadtree Segmentation

The main step of our method is the quadtree segmentation part, which is described in this section. A *quadtree* scanning of the image is the core step of the segmentation. The image is divided into four regions, and each of these regions is compared with their adjacent 4-neighbors using a comparison operator. If two regions are evaluated as similar they are merged. Regions that are not merged with any other region are divided in four new regions and the same comparison operation with their new neighbors is done. This process is performed recursively until there are no more regions to divide or the region size has reached the *mq*s parameter.

The comparison operator takes two adjacent regions and considers the *Constraints Image* built in step 2. It counts pixel borders on each region and use predefined rules to check if there is a border

compatibility between these regions. If no border compatibility is found the operator evaluates them as non similar, otherwise it constructs color histograms for each region and then normalize and eliminate noise in them. Next, it calculates the Euclidean distance between these histograms. If this distance is less than the  $ht$  parameter, the regions are evaluated as similar. Regions merged during this step give the shape of the segmentation result. Regions that could not be assigned to any segment due to absence of homogeneity or similarity are ignored and left out of the final segmentation result. This behavior helps our algorithm to produce more homogeneous (noise free) segments.

Due to the progressive approach of this segmentation procedure, we can take advantage of the parameter  $mqs$  to shorten the segmentation time according to our needs, although we will always have an approximation to the final segmentation result. This is a desirable property of segmentation methods for time restricted applications such as video processing or CBIR. Figure 3 depicts the quadtree processing of a sample image, it also shows how can different values for the  $mqs$  parameter support anytime segmentation.

### 3.5 Segmentation Enhancement

The final step of the proposed methodology is the enhancement of segmentation result described in last section. Segmentation enhancement attempts to improve segmentation results in two steps. First, it applies the comparison operator to all segments that intersected during step 4 and merge them accordingly. Second, it looks for segments that do not meet the  $mos$  parameter and merges them with the most similar adjacent segment that does. Figure 4 shows an example of image segmentation using our proposal.

### 3.6 Complexity analysis

The complexity of the whole process of our proposal is  $O(N) = N \times \sqrt{N}$ , where  $N = w \times h$  is the number of pixels of the image. Where the complexity of each step of the proposed methodology is as follows: (1)  $7 \times N$ ; (2)  $4N$ ; (3)  $N$ ; (4)  $2^{\log_4(N-1)}6N + C$ , with  $C$  the number of operations required to divide a quadtree; (5)  $2K + K^2 + N$  with  $K$  the number of segments found. A segmentation algorithm with complexity of  $O(N) = N \times \sqrt{N}$  can be considered efficient as most segmentation algorithms are of order  $O(N^2)$  and higher [5]. Further, one should note that the proposed method is an *any-time* technique.

## 4 Experimental settings

This section describes the experimental setting we adopted for the evaluation of the proposed segmentation method. The goal of our experiments is: 1) to evaluate the segmentation performance of the proposed method and 2) to evaluate the performance of an annotation method with segments generated with our segmentation technique in terms of segmentation and annotation accuracy.

For the evaluation of the segmentation performance and efficiency we considered three datasets of images. The first one consists of 18 heterogeneous manually selected images; the second one consists of 100 randomly selected images from the Berkeley Segmentation Dataset [14]; and the third dataset consists of 14 images randomly selected from “Image of the Day” Wikipedia section [15]. These datasets were used to progressively enhance the algorithm attempting to produce acceptable weak segmentation results for more images as possible in all the three datasets. The parameters of the segmentation algorithm were fixed by trial and error. The values of those parameters are  $mos = 1.25\%$ ;  $mqs = 2$ ;  $ht = 0.5$ , with these values we have obtained acceptable results for the different datasets, hence they can be considered default values.

The annotation performance was evaluated with a dataset consisting of 500 images taken from the SAIAPR TC-12 [4], which is a benchmark of manually segmented and annotated images. We compare

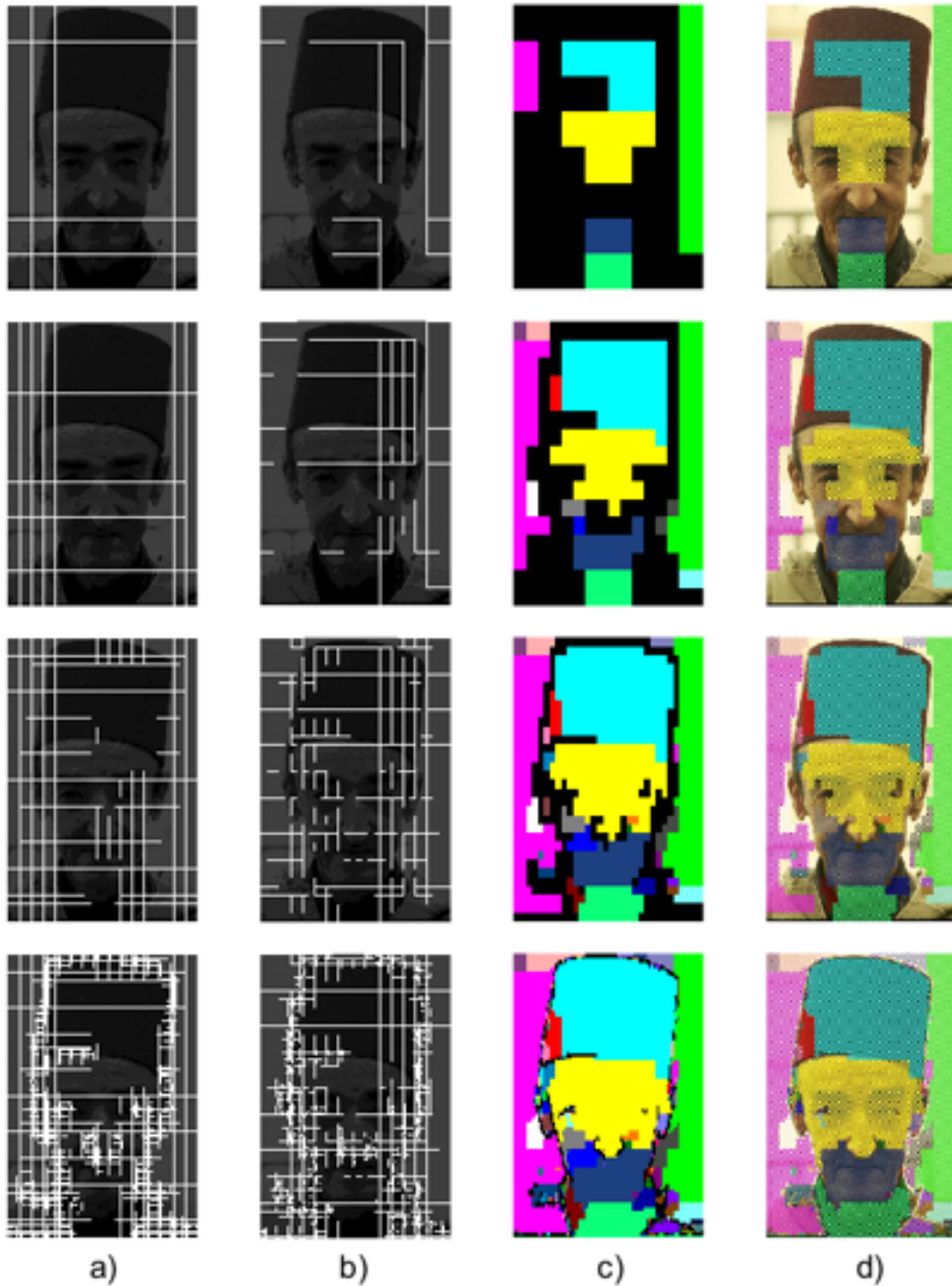


Figure 3: Segmentation results using our proposal with different values for the  $mos$  parameter. On first row  $mos$  is 32, on second row is 16, on third row is 4 and on fourth row is 2. Column *a*) shows all the generated quadtree regions, column *b* shows only the merged regions, column *c* shows colored results and column *d* shows the overlapping of the segmentation result on the original image. The progressive nature of our proposal makes possible to have an approximation to the final result on each step.

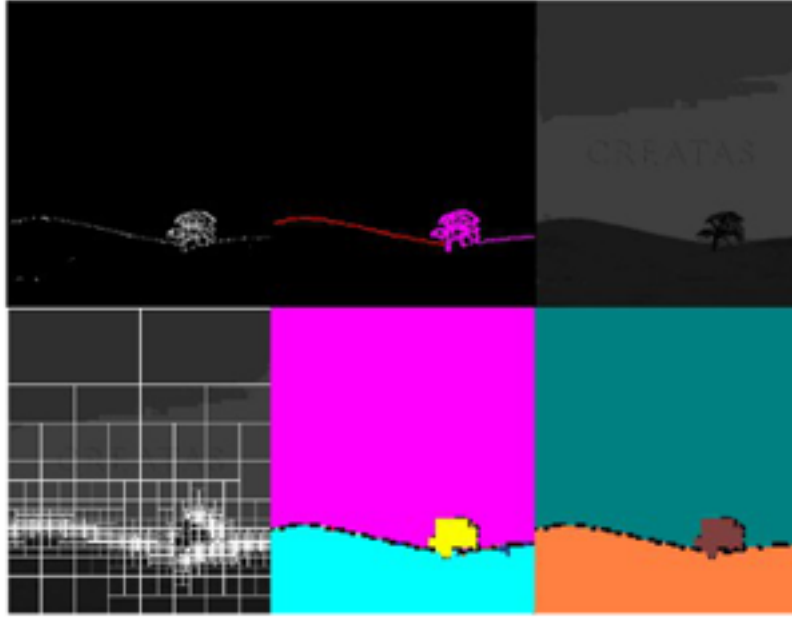


Figure 4: Visual results of each step of our segmentation proposal. From right to left and from top to down we have: border detection, border selection, color discretization, quadtree segmentation process, quadtree segmentation colored results, enhanced segmentation after post-processing.

our proposal with two different segmentation algorithms used in state of the art image annotation: *normalized cuts* and *grid segmentation*. Normalized cuts consists of building a weighted graph from the image in which each node represents a pixel and the arcs' values are the similarities between connected pixels using a function that measures more than 30 features on each pixel [5]. Grid segmentation consists of cutting the image by grid of  $n$  rows and  $m$  columns of equal size. After segmenting the 500 images dataset with the three mentioned algorithms, segmentation results from each segmentation algorithm were given to an external automated annotation engine [3]. Then, we compared the annotations obtained for each image with each of the segmentation algorithms.

The evaluation process for image annotation is as follows. For each set of segmentation results given by each segmentation algorithm for the 500 images, we take each image and match their automatically produced segments with the corresponding manually produced segments. Then we check if the annotation in both segments is the same; more specifically, we estimate the percentage of pixels that were correctly labeled. Intuitively, images receive points for each correct annotation, maximum grade for a perfectly annotated image is 1. The points given for each correct annotation are weighted according to segment's size. This means that if an image has 3 annotations in its manually annotated version, and one annotation corresponds to a segment that occupies 75% of the total image area, then that segment will give 0.75 points in case it is correctly annotated. The execution time that spent each algorithm to segment the 500 images was also recorded.

Our testing environment was an AMD Turion 64 bits 1.6GHz 1GB RAM laptop running Windows XP. The reason for this choice was that we wanted to assure the performance of our proposal within a common and accessible environment.

## 5 Results

Sample segmentation results obtained with the proposed method are shown Figure 5. From this figure we can see that segmentations produced with the proposed method are more accurate than those generated with normalized cuts. It is common for normalized cuts to partition regions that belong to the same object in more than a single region, the proposed method, on the other hand, correctly identifies homogeneous segments. One should note that, as stated above, segmentation is a highly subjective process, hence evaluating the quality of segmentations produced with different methods is a subjective task as well. Nevertheless, we believe the segmentations produced with our method provide very useful information about the objects present in the image.

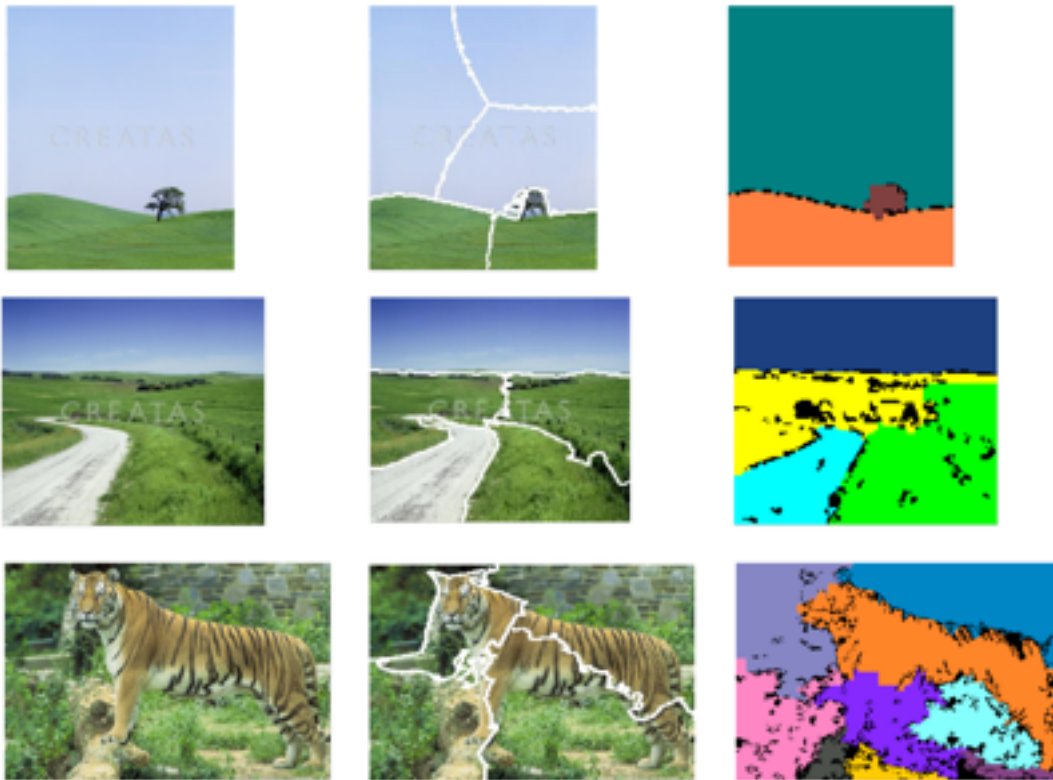


Figure 5: Example of the segmentation results obtained with normalized cuts (middle) and with the proposed quadtree segmentation algorithm (right).

Table 1 summarizes the experimental results obtained with the proposed algorithm in terms of both image segmentation and image annotation performance. Sample annotation results are shown in Figures 6 and 7.

Criterion	Quadtree	Normalized Cuts	Grid Segmentation
Seconds required to segment 500 images	<b>1,126.83</b>	20,000.00	74.40
Total annotation precision of 500 images	<b>30.8258</b>	22.7632	20.3786
Per image annotation average	<b>0.06177</b>	0.04570	0.04100
Standard deviation	<b>0.08465</b>	0.06054	0.07216

Table 1: Results for segmentation time and annotation precision of our experimental evaluation.



From Table 1 we can see that the *Quadtree Segmentation Algorithm* produced segments that were annotated 50% better than those produced with *grid segmentation* and 35% better than those generated with *normalized cuts*. In addition, our proposal segments about 17 times faster than *normalized cuts*. Despite the fact that the proposed method was less efficient than the grid approach, the 50% improvement makes worthwhile using our method. Whereas the improvement over the grid approach is somewhat expected, the improvement over the normalized cuts technique is an important result, as normalized cuts has been the most used image segmentation algorithm in the context of automatic image annotation [1, 2, 3, 11, 16].

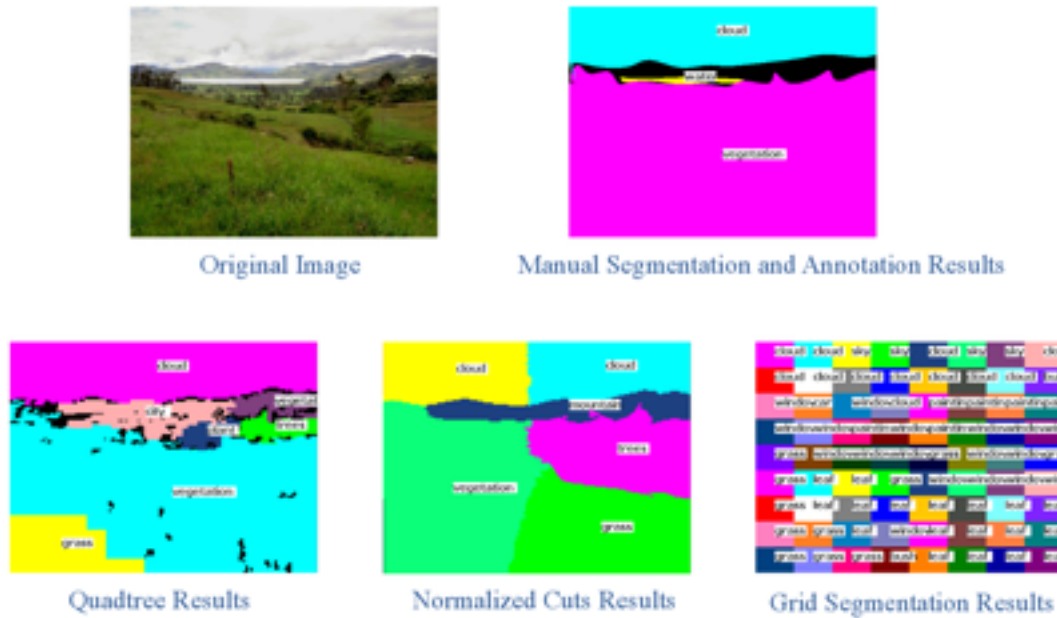


Figure 6: Example 1 of the segmentation and annotation results obtained with the proposed quadtree segmentation algorithm.

Figures 6 and 7 confirm that the proposed method provides segments that are better classified than those produced with the other techniques. Nevertheless, we must emphasize that the proposed algorithm is somewhat dependent on the homogeneity and borders of images; that is, when images lack homogeneity or do not present well defined borders the method the precision of the proposed method is reduced. In future work we will explore the use of other features than can overcome the drawbacks of the current implementation.

## 6 Conclusions

We introduced a simple and efficient method for image segmentation, which resulted very helpful for image annotation, as evidenced by experimental results. This is specially true for images with big objects and simple textures. The *any time* segmentation feature of our proposal in combination with its order of complexity makes it attractive to be used in time restricted environments. For the experiments reported in this paper, parameters were set by trial an error, it would be interesting to develop methods that automatically can tune parameters according the type of images. Future work also includes improving the selection of relevance borders and extending the color and texture features used in the comparison operator, this in order to improve the segmentation results in particular for complex images. The multilevel annotation taking advantage of the hierarchical nature of quadtree is also encouraging. In addition, a parallel implementation of the algorithm can be done in order speed even more our formulation.

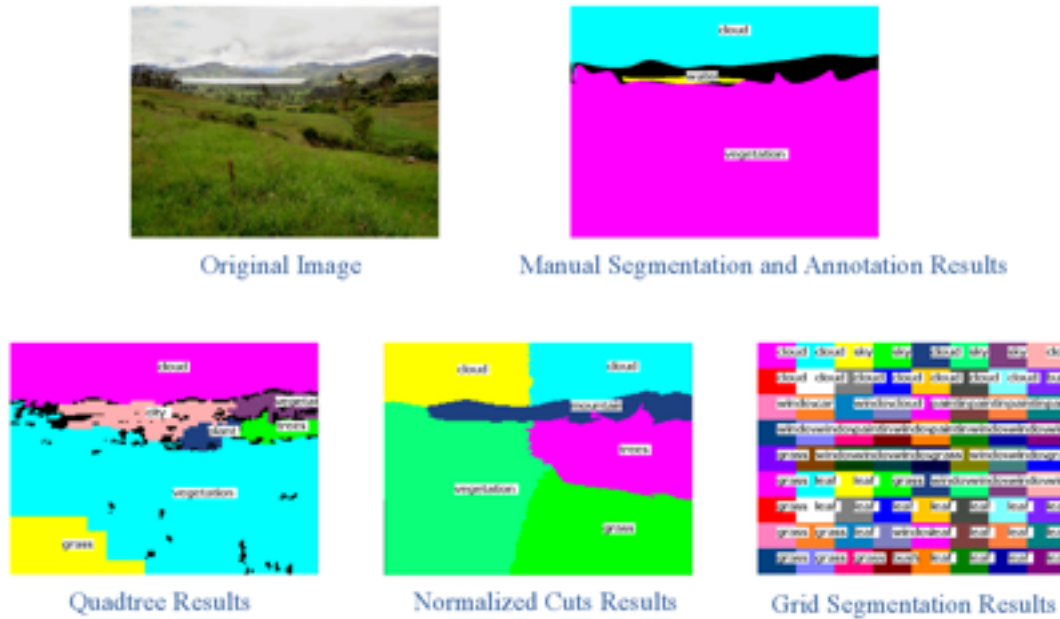


Figure 7: Example 2 of the segmentation and annotation results obtained with the proposed quadtree segmentation algorithm.

## References

- [1] K. Barnard, Q. Fan, R. Swaminathan, A. Hoogs, R. Collins, P. Rondot, and J. Kaufhold. “Evaluation of Localized Semantics: Data, Methodology, and Experiments”, *International Journal of Computer Vision*, Vol. 77 (1-3), pp. 199–217, 2008.
- [2] H. J. Escalante, M. Montes, and L. E. Sucar. “Word Co-occurrence and Markov Random Fields for Improving Automatic Image Annotation”, *Proceedings of the 18th British Machine Vision Conference*, Vol. 2, pp. 600–609, Warwick, UK, 2007.
- [3] H. J. Escalante, M. Montes and L. E. Sucar. “An Energy-based Model for Region Labeling”, *Computer Vision and Image Understanding*, in press, 2011.
- [4] H. J. Escalante, M. Grubinger, C. A. Hernández, J. A. González, A. López, M. Montes, E. Morales, L. E. Sucar, and L. Villaseñor. “The Segmented and Annotated IAPR TC-12 Benchmark”, *Computer Vision and Image Understanding*, Vol. 114, pp. 419–428, 2010.
- [5] J. Shi and J. Malik. “Normalized Cuts and Image Segmentation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22 (8), pp. 888–905, 2000.
- [6] L. T. Lan and A. Boucher. “An Interactive Image Retrieval System: from Symbolic to Semantic” *Proceedings of the International Conference on Electronics, Informations and Communications*, Hanoi, Vietnam, 2004.
- [7] G. Papadopoulos, V. Mezaris, I. Kompatsiaris, and M. Strintzis. “Combining Global and Local Information for Knowledge-Assisted Image Analysis and Classification”, *EURASIP Journal on Advances in Signal Processing*, Vol. 2007, Article ID 45842, 15 pages, 2007.
- [8] C. Saathoff and S. Staab. “Exploiting Spatial Context in Image Region Labelling using Fuzzy Constraint Reasoning”, *Proceedings of the 9th International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 16–19, IEEE, Klagenfurt, Austria, 2008.
- [9] D. A. Forsyth and J. Ponce. “Computer Vision a Modern Approach”, *Prentice Hall*, 2002.
- [10] T. Pavlidis. “Segmentation of Pictures and Maps through Functional Approximation”, *Computer Graphics and Image Processing*, Vol. 1 (4), pp. 360–372, 1972.
- [11] K. Barnard, P. Duygulu, N. de Freitas, D. A. Forsyth, D. Blei, and M. I. Jordan. “Matching Words and Pictures”, *Journal of Machine Learning Research*, Vol. 3, pp. 1107–1135, 2003.

- [12] C. Galleguillos, and S. Belongie. “Context Based Object Categorization: A Critical Survey”, *Computer Vision and Image Understanding*, (CVIU), Vol. 114, pp. 712–722, 2010.
- [13] H. Samet. “The Quadtree and Related Hierarchical Data Structures”, *ACM Computing Surveys*, Vol. 16 (2), pp. 187–260, 1984.
- [14] <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>
- [15] [http://en.wikipedia.org/wiki/Wikipedia:Picture\\_of\\_the\\_day/Archive/](http://en.wikipedia.org/wiki/Wikipedia:Picture_of_the_day/Archive/)
- [16] P. Carbonetto, N. de Freitas, and K. Barnard. “A Statistical Model for General Context Object Recognition”, *Proceedings of the 8th European Conference on Computer Vision*, LNCS Vol. 3021, pp. 350–362, Springer, Prague, Czech Republic, 2004.