Fast Two-Step Segmentation of Natural Color Scenes using Hierarchical Region-Growing and a Color-Gradient Network

Aldo v. Wangenheim[1], Rafael F. Bertoldi[1], Daniel D. Abdala[2], Michael M. Richter[3], Lutz Priese[4], Frank Schmitt[4]

[1]Image Proc. and Computer Graphics Lab – LAPIX, Federal University of Santa Catarina, Florianópolis, Brazil {awangenh,fogo}@inf.ufsc.br

[2]CS Department, University of Münster, Münster, Germany abdalad@uni-muenster.de

[3]CS Department. University of Calgary, Alberta, Canada mrichter@cpsc.ucalgary.ca

[4]CS Department, University of Koblenz, Koblenz, Germany {priese,fschmitt@uni-koblenz.de

Abstract

This paper presents and empirically evaluates a new combined approach for the fast segmentation of complex color images, in particular images presenting color structures with strong but continuous color or luminosity changes, such as commonly found in outdoors scenes. The approach is based on the combination of a pre-segmentation based on a fast hierarchical method, used in combination with an efficient post-segmentation based on the graph-based analysis of global color and luminosity gradients. This combination of two different and fast region-growing segmentation methods overcomes problems shown by both methods when used alone. We also show empirically that the quality of the segmentations generated by this two-step approach is very promising and comparable to segmentations generated by state-of-the-art segmentation methods that were available for comparison when this paper was being written.

Keywords: color image segmentation, fast segmentation, outdoor scenes, Color Structure Code, Gradient Network Segmentation

1. Introduction

Natural color scenes, such as outdoors images composed by many colored objects that are acquired under uncontrolled conditions show complex illumination patterns across the same object in the picture. Examples are variations in lightness and specular effects. State-of-the-art region-growing segmentation methods [2][3][10] present two main features that limit their applicability for dealing efficiently with natural scenes:

- **A static region similarity concept**, where pixels or textures within a region are expected to be homogeneous. Typical natural scenes, however, show strong continuous variations of color, presenting a different, dynamic order that is not taken into account by these algorithms. They will e.g. segment a sky region with different intensities of blue or will represent an irregularly illuminated surface as a set of different regions. When the parameters of such algorithms are stressed in order to try to accomplish a correct segmentation of a large object showing a long continuous gradient of color, typically with a gradual but large color variation, a region leakage of other objects in the image is likely to occur. Then the algorithm is becoming unstable and even inapplicable.
- **Increase in complexity** to present more stable results, that usually demand complex computations to detect segment-correlation clues or they are built upon additional texture information. This slows down considerably the processing time without being much more stable when extreme color variations are present.

In this paper we present an improvement of our previsously presented gradient network approach for image segmentation [1] and validate it empirically in conjunction with another method [2], providing a novel color scene segmentation approach that is extremely fast, providing a two-step approach that shows satisfactory results when applied to natural scenes, while not showing poorer performance than state-of-the-art methods when applied to standard region-growing problems. The approach described here is based on a pipeline of two fast segmentation algorithms: The first step is a hierarchical region-growing segmentation [2] that generates an oversegmented picture where natural boundaries of objects are preserved.

The second step is based on a color gradient based region-growing segmentation method [1]. It starts from the initial segmentation rather than from a pixel image and computes a gradient network that spans the entire image and finds locally connected gradients that show an organized pattern, representing ordered color variations of a same object. These “organized segment clusters” composed of different shades of color and light intensity that correlate to each other are melted into meta-regions that are then presented as the final result of the segmentation. One of the most important features of this algorithm is its computational efficiency.

In this paper we will present a short review of the *Gradient Network Method* (GNM) with special attention to the features that are responsible for its speed. The *Color Structure Code* (CSC) method will also be addressed, followed by results obtained with this combined approach. A processing time comparison is presented taking into account different combinations of region growing algorithms with the GNM. It should clarify the main reason for choosing the CSC as the first stage of this approach. Finally we present a discussion that takes into account the applicability of such an algorithm for real-time color image segmentation.

2. Gradient Network Method

The *Gradient Network Method* [1] was developed to deal with segmentation problems where objects in the scene will be represented by several different but similar and gradually varying color shades, as they often are found in outdoors scenes. Besides being a technique that uses a novel segmentation strategy, GNM is also a technique that shows good results in terms of efficiency.

The GNM looks for a higher degree of organization in the structure of the scene through search and identification of continuous and smooth color gradients. To be able to run over the image and identify these variations of colors, the GNM uses a graph *G (V, E)* to structure the initial stage of the algorithm. The graph will be used as a structure to guide the algorithm. This strategy is related to the approach in [12]. The vertices $v \in V$ represent regions identified in a previous pre-segmentation. GNM concentrates on regions of high similarity, specifically in the aspect of low color variation. The goal of the pre-segmentation with a different algorithm is to obtain groups of pixels with a high degree of similarity represented in a simple way, avoiding possible problems with local noise if the representation would be done individually for each pixel as a vertex.

The tests with GNM described here were performed combining it with a Mumford-Shah functional based segmentation provided by the Megawave package [16] and with the *Color Structured Code* (CSC) [2] algorithm. Here we remark that techniques like Watershed [13] or similar ones could also be used. The choice of the particular algorithm relies basically in it fitness in producing super-segmented results that preserve the main edges. The external pre-segmentation step is followed by a labeling procedure to convert the segmentation output into a graph *G (V, E)*.

After creating the graph *G(V,E)*, the next step is checking all the neighborhood relations if they comply to the similarity measure and provide continuous and smooth color gradients. The evaluation of the continuity of the gradients along the paths found in the graph is done by a function f_c that takes into account the *perception* [6] variations. This allows a better evaluation of the similarity in presence of different luminances in the regions. Most long continuous and smooth gradients are results of the presence of lighting effects in the scene of an image. With this additional feature, the algorithm becomes more robust for such characteristics. Therefore, even when the neighborhood contains regions too dark or too illuminated it will search for the best possible gradient path in the graph [1].

All $e \in E$ will be evaluated by the chosen similarity measure and regions found acceptably similar will be grouped in meta-regions. The resulting meta-regions of the whole process will be the output produced by the GNM segmentation. A general description of the algorithm can be found in figure 1.

3. The Color Structure Code

The Color Structure Code (CSC) [2] was developed at the CS Department of the University of Koblenz, Germany. CSC was aimed at the segmentation of scenes from a camera in a car in motion for realtime road sign recognition. The CSC is a region growing algorithm that uses a hierarchical topology formed by islands, a topology type introduced by [24]. These islands have different levels. For instance, a level 0 island is a hexagon, composed by the 6 vertex points around the central point. During the process, some islands overlap others in a way that level n+1 islands are composed by seven level n overlapped islands. This will be repeated until an island covers the entire given image.

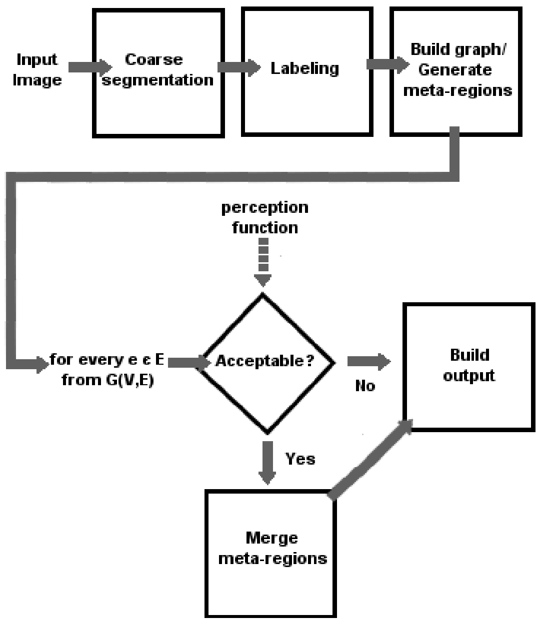


Figure 1. Diagram displaying the GNM process

As a first step, the whole image will be partitioned into level 0 islands. A merging step will follow where the islands will grow and overlap iteratively. Next to the grouping step is the split step, where some corrections will take place through the use of global information.

In this way, CSC combines a local information step in the merging process and a global information evaluation in the split step, looking for segmenting regions with the highest similarity degree.

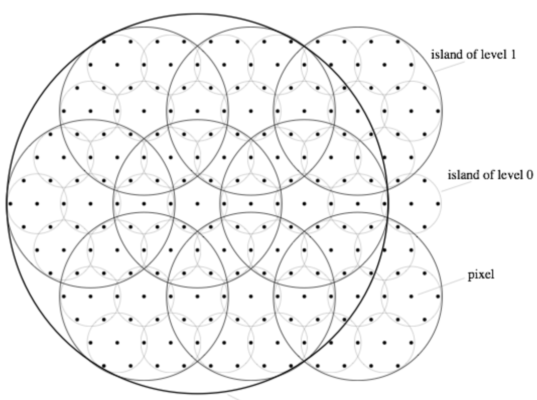


Figure 2. CSC's hierarchical island structure [25]

Regarding outdoor images, CSC alone is already a good alternative to deal with images of this kind. However, as it is not prepared to deal with long and slowly varying color regions, segmentation problems will occur. As a usual consequence found in most algorithms, sensitive regions might be swallowed or more cautious parameters might produce many more segments than expected. We will show that the CSC shows good performance and reliability regarding outdoor images when combined with GNM.

4. Achieving Results Combining Efficiency and Image Quality

High speed performance segmentation algorithms have been investigated for satisfying the demand from applications that require real-time or almost real time results. Fast segmentation processes could be used in several situations, like motion detection in video frames or autonomous vehicle guidance [2]. Another application would be to guide surgery and other medical procedures. An example is given in [17], where segmenting a carotid artery is a useful step in medical imaging. Efficiency requirements can also be found in several other areas. [18] presents a technique developed for real time applications as space-weather analysis. In [19] an approach is proposed to track players in a soccer pitch. Other applications can be found in the detection of defects in production lines [20]. Fast segmentation approaches are a recurrent topic and several optimizations or specializations over known techniques have been developed [21][22][23].

While several algorithms can achieve good results neglecting speed, GNM and CSC are both generic segmentation techniques that provide a reasonably good performance.

4.1. GNM Complexity Issues

GNM achieves performance through a set of integrated strategies. First, an optimized labeling algorithm performs the initial processing of the pre-processed image and ensures a fast solution to this intermediate step. The complexity of the used labeling algorithm is $O(n^2)$.

After classifying the information in the labeling, the construction of the graph takes place. This will structure the information since every region found by the labeling will correspond to a vertex of the graph. The graph generation step has a complexity of $O(n)$. To improve the performance and avoiding redundant loops, the mean color value computation for every region and the conversion to the HSI color space [26] are done together with the graph generation step.

When the graph is built, it is traversed for searching perceptibly similar regions that should be merged. Since this step depends solely on the

Figure 3 - CSC results when dealing with a complex outdoors image. From left to right columns, respectively: the original image, a cautious approach and an aggressive approach.

number of edges, its complexity is *O(m)*.

GNM total complexity is $O(n^2 + n + m)$, where n is the number of vertices and m the number of edges. This method presents a simple solution that is only dependent of the image size and the scene complexity of the resulting pre-processed image. It is important to note, though it can't be accounted in the GNM complexity, that the chosen algorithm for the pre-segmentation has an effect on the total time of processing in this approach. A proper technique must be selected here.

4.2. Pre-Segmentation Issues

As our main focus here is to obtain results of good quality combined with high speed performance, we have chosen the *Color Structure Code* (CSC) [2] as our pre-segmentation technique.

Though CSC is focused on speed and was developed for specific purposes, it still achieves good results in terms of image quality and proves to be a good solution in generic cases too. The islands of similarity approach fits nicely with the expected feature for GNM starting point, *i.e.* the regions of very similar characteristics avoiding leakages.

As a main source for comparison, we used the Mumford-Shah Functional implementation supplied by the *Megawave* [16] image processing package. The behaviour of this method is well known and documented and was considered the best choice for performance and quality comparisons.

5. Results and Discussion

To empirically validate the approach, several outdoors images where comparatively processed with a set of different algorithms. To allow comparisons, all tests were performed on the same computer with the algorithms presented below.

The adoption of the Berkeley's image dataset [8] was a necessary and desirable choice since it is a well known dataset with the added features of ground truth (hand segmented) images for every set that will help us in future quality evaluations. The comparison with other currently developed techniques and results are easier obtained by the readers since this dataset is used by several other authors and publications.
The segmentation techniques used in the following tests were: a) GNM applied over pre-segmented images by CSC, with a threshold equal to 30 (total time will the sum of both algorithms execution); b) GNM applied over pre-segmented images by a *Mumford-Shah* functional based segmentation, with lambda equal to 600 (total time will the sum of both algorithms execution); c) CSC, with a threshold selected to produce the best achievable result; d) *Mumford-Shah* functional based segmentation, with a *lambda* parameter selected to produce a result considered the best possible; e) JSEG [3], a well-known unsupervised segmentation technique.

Image	CSC+GNM	M&S+GNM	CSC	M&S	JSEG
2092	0.844s (0.641)	5.453s (0.531)	0.171s	4.782s	9.297s
14037	0.89s (0.64)	9.265s (0.578)	0.203s	8.172s	11.141s
15088	0.938s (0.688)	5.627s (0.625)	0.203s	5.063s	26.5s
24044	1.000s (0.734)	5.719s (0.719)	0.234s	4.969s	10.78s1
24063	0.797s (0.594)	5.25s (0.5)	0.281s	4.828s	9.078s
46076	0.828s (0.641)	5.36s (0.547)	0.172s	4.766s	11.094s
67079	1.015s (0.656)	5.437s (0.593)	0.235s	4.937s	12.328s
295087	1.015s (0.781)	5.171s (0.578)	0.188s	4.437s	14.156s
310007	0.828s (0.625)	5.641s (0.532)	0.188s	5.031s	9.89s
368078	0.969s (0.735)	5.156s (0.656)	0.172s	4.547s	12.578s
3096	0.625s (0.547)	5.25s (0.484)	0.063s	4.641s	7.485s
22090	0.687s (0.593)	5.125s (0.469)	0.078s	4.687s	9.063s
48055	1.077s (0.968)	5.454s (0.532)	0.094s	4.828s	14.219s
138078	0.75s (0.656)	5.968s (0.5)	0.078s	5.36s	14.359s
124084	0.828s (0.734)	5.016s (0.469)	0.078s	5.375s	9.39s
118035	0.749s (0.64)	6.03s (0.468)	0.078s	5.375s	9.39s
143090	0.781s (0.687)	4.938s (0.469)	0.078s	4.469s	12.907s
219090	0.766s (0.672)	5.234s (0.468)	0.093s	4.735s	11.844s
Mean	**0.855s (0.68)**	**5.616s (0.54)**	**0.149s**	**5.056s**	**11.972s**
Standard deviation	**0.125s (0.092)**	**0.960s (0.073)**	**0.069s**	**0.829s**	**4.149s**

Table 1. The total execution time for a set of 18 Berkeley dataset images processed using the techniques is listed in the Results and discussion session. In the GNM cases, the time consumed by GNM is shown inside of the parenthesis. Landscape images dimensions are 481x321 pixels and portrait images are 321x481 pixels. The mean time for every algorithm test is displayed in the last line of the table.

The total execution time for each set with every selected algorithm is shown in **Erro! Fonte de referência não encontrada.**. This time was obtained by the difference of two time stamps, one in the start and one in the end of the execution process of each algorithm. Mean and standard deviation for every set are also displayed.

The computer the tests were run on is an AMD Athlon 64, 2.2 GHz with 512MB RAM memory and the time unit is seconds. **Erro! Fonte de referência não encontrada.** shows image results obtained with GNM combined with both CSC and Mumford-Shah. As the **Erro! Fonte de referência não encontrada.** shows, the combination of CSC and GNM shows results with a mean value of less than 1 second, which is slower than CSC alone. This was expected, considering the cumulative times of CSC+GNM. The mean time is several times lower than Mumford-Shah and JSEG. There is little deviation among the times obtained for CSC+GNM, while again in accordance with the exception of CSC alone, all other techniques show higher deviations.

Comparing CSC+GNM with MS+GNM, we see that GNM takes longer in the CSC+GNM case than in the MS+GNM case. This occurs because of the existence of little fragments that are produced by CSC which are not found in Mumford-Shah segmentations, resulting in much more graph vertices to be evaluated. It is important to notice, however, than GNM has a very stable performance in both cases, with little deviation among the tests cases.

Comparatively, CSC+GNM presented a stable performance while providing robust image results. Table 2 shows segmentations of complex illuminated objects, as the sky in the 368078 set or the red roof of the church in the 118035 set. Higher resolution images, comparisons among several algorithms and more results can be found in www.lapix.ufsc.br/fast.

6. Conclusions

We have empirically shown that the quality of the segmentations generated by our two-step approach is very promising and comparable to segmentations generated by state-of-the-art segmentation methods that were available for comparison when this paper was being written.

On the other side, segmentation time of a given image when processed by our suggested two-step method was shown to be considerably less than when other approaches were used or when the Gradient Network step was used in combination with more traditional segmentation approaches such as the Mumford-Shah functional.

The Gradient Network Method is a segmentation post-processing method that is independent of the region-growing method that is applied to generate the super-segmented input image. This has been shown by the comparison between the results produced using the CSC method and when the Mumford-Shah functional is used as the pre-processing step. It is important to note that the quality of the final results are very similar, although the intermediate segmentation results of the Mumford-Shah functional are sometimes of a "prettier" quality. The processing time, however, is extremely shorter when a rapid approach like the CSC, which was originally developed for real-time color segmentation, is used. This shows that the processing step with the Gradient Network Method allows us to rely on very fast pre-segmentation methods that reduce the total processing time while producing end-segmentations of good quality.

Further improvements could still be achieved in terms of efficiency with the use of a graphics processing unit for performing the necessary computations of the involved algorithms. This kind of technology, referred as General-Purpose Computing on Graphics Processing Units (GPGPU), would achieve better results, probably real-time ones. This could make the combination of CSC and GNM a feasible solution to real-time applications that deal with outdoor scenes, as robotics or traffic monitoring applications. Preliminary results not reported and shown here gave some promising perspectives..

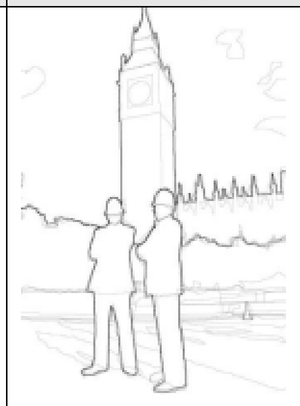
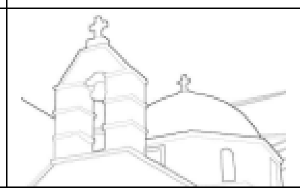
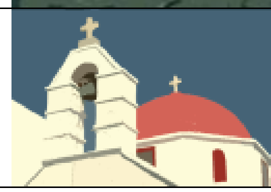
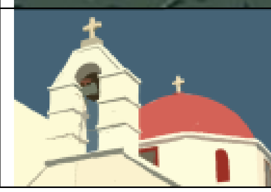
Original	Ground-truth	CSC+GNM	M&S+GNM
			
			

Table 2. Examples of the image results obtained by GNM, combined with both CSC and Mumford-Shah. The first row corresponds to image number 368078 and the second row corresponds to 118035 from Berkeley image dataset. The ground-truth image is also provided by Berkeley dataset. More sets, results comparing to several techniques and higher resolution images can be found in www.lapix.ufsc.br/fast. We also refer to [1] to view some image comparisons.

Acknowledgements

Daniel D. Abdala thanks the CNPq-Brazil for a Ph.D scholarship, and Professor Xiaoyi Jiang for his support during the first months of the Ph.D. work at the University of Münster.

References

[1] Wangenheim A, Bertoldi R, Abdala D, Richter M. Color image segmentation guided by a color gradient network. doi:10.1016/j.patrec.2007.05.009.

[2] Rehrmann, V. ,Priese, L.: Fast and Robust Segmentation of Natural Color Scenes. ACCV (1), 1998: 598-606

[3] Deng Y, Manjunath BS. Unsupervised segmentation of color-texture regions in images and video. IEEE Transactions on Pattern Analysis and Machine Intelligence; 2001;23(8):800-810.

[4] Dony, R.D., Wesolkowski, S.B. Edge Detection on Color Images Using RGB Vector Angles. Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering. Edmonton, Alberta, Canada May 9-12 1999, ,pg. 687-692.

[5] Dupuis, A. and Vasseur, P. Image

segmentation by cue selection and integration, Image and Vision Computing, Volume 24, Issue 10, October 2006: 1053-1064.

[6] Huang K, Wang Q, Wu Z. Natural color image enhancement and evaluation algorithm based on human visual system. Computer Vision and Image Understanding; July 2006; 103(1):52-63.

[7] Kato, Z. and Pong, T.-C., A Markov random field image segmentation model for color textured images, Image and Vision Computing, Volume 24, Issue 10, October 2006: 1103-1114.

[8] Martin D, Fowlkes C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proc. 8th Int'l Conf. Computer Vision. vol. 2; 2001. p. 416-423.

[9] Mittelheußer, Gangolf und Kruggel, Frithjof: Fast Segmentation of Brain Magnetic Resonance Tomograms. VRMed‘95: First International Conference on Computer Vision, Virtual reality and Robotics in Medicine, INRIA, Nizza, Springer Verlag, April, 1995.

[10] Mumford, D. and Shah, J. Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math., vol. 42, 1989: 577–684.

[11] Schneider, M.K.; Fieguth, P.W.; Karl, W.C.; Willsky, A.S.; Multiscale methods for the segmentation and reconstruction of signals and images. IEEE Transactions on Image Processing, Volume 9, Issue 3, March 2000: 456 - 468.

[12] Trémeau A, Colantoni P. Regions adjacency graph applied to color image segmentation. IEEE Transactions on Image Processing; 2000;9(4):735-744.

[13] Vincent, L. and Soille,P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Analysis and Machine Intell., vol. 13, pp. 583--598, 1991.

[14] Wesolkowski, S.B. Color Image Edge Detection and Segmentation: A Comparison of the Vector Angle and the Euclidean Distance Color Similarity Measures. Master Thesis. University of Waterloo, Ontario, Canada, 1999.

[15] Wu, Z.; Leahy, R.; An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 15, Issue 11, Nov. 1993:1101 – 1113.

[16] Megawave image processing package, accessible through http://www.cmla.ens-cachan.fr/Cmla/Megawave/. Access in: 14 September 2006.

[17] Kim D Y, Park J W. Connectivity-based local adaptive thresholding for carotid. Image and Vision Computing; 2005; 23(14): 1277-1287.

[18] de Wit T D. Fast Segmentation of Solar Extreme Ultraviolet Images. Solar Physics; 2006; 239(1-2):519-530.

[19] Figueroa P, Leite N J, Barros R. Background recovering in outdoor image sequences: an example of soccer players segmentation. Image and Vision Computing; 2006; 24(4): 363-374.

[20] Abouelela A, Abbas H M, Eldeeb H, Wahdan A A, Nassar S M. Automated vision system for localizing structural defects in textile fabrics. Pattern Recognition Letters; 2005; 26(10): 1435-1443.

[21] Pan Y, Birdwell J D, Djouadi S. Efficient Implementation of the Chan-Vese Models Without Solving PDEs. Multimedia Signal Processing 2006 IEEE 8th Workshop; 350-354.

[22] Wong K Y, Spetsakis M E. Tracking based motion segmentation under relaxed statistical assumptions. Computer Vision and Image Understanding; 2006; 101(1); 45-64.

[23] Sun H, Yang J, Ren M. A fast watershed algorithm based on chain code and its application in image segmentation. Pattern Recognition Letters; 2005; 26(9); 1266-1274.

[24] Hartmann, G. Recognition of hierarchically encoded images by technical recognition of hierarchically encoded images by technical and biological systems. Biological Cybernetics, v. 57, n. 1–2, p. 73–84, August 1987.

[25] Priese, L.; Sturm, P. Introduction to the color structure code and its implementation. March 2003. Accessible through: <http://www.uni-koblenz.de/lb/lb_downloads/download-/csc.pdf>.

[26] Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang , Jingli. Color image segmentation: advances and prospects. Pattern Recognition; 2001; 34; 2259-2281.