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

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

[1]Lab for Image Processing and Computer Graphics – LAPIX
Department of Computer Sciences – INE, Universidade Federal de Santa Catarina – UFSC
Phone: +55 (48) 3721-9516 (LAPIX), +55 (48) 3721-9117 (Telemedicine)
CEP 88049-200 Florianopolis - SC, Brazil
FAX: +55 (48) 3331-9516-R16
{awangenh | fogo}@inf.ufsc.br

[2]University of Münster, Department of Computer Science
Einsteinstrasse 62, D-48149 Münster, Germany
Phone: +49 (251) 833-3759, Fax: +49 (251) 833-3755
abdalad@uni-muenster.de

[3]Department of Computer Science, University of Calgary
Phone: +1 (403) 220-6783, Fax: +1 (403) 284-4707
Calgary, Canada
mrichter@ucalgary.ca

[4]Institut für Computervisualistik, Universität Koblenz-Landau
Phone: +49 (261) 287-2792, Fax: +49 (261)287-100-2792
Postfach 201 602, 56016 Koblenz, Germany
{priese | fschmitt}@uni-koblenz.de

Abstract

We present evaluation results with focus on combined image and efficiency performance of the Gradient Network Method to segment color images, especially images showing outdoor scenes. A brief review of the techniques, Gradient Network Method and Color Structure Code, is also presented. Different region-growing segmentation results are compared against ground truth images using segmentation evaluation indices Rand and Bipartite Graph Matching. These results are also confronted with other well established segmentation methods (EDISON and JSEG). Our preliminary results show reasonable performance in comparison to several state-of-art segmentation techniques, while also showing very promising results comparatively in the terms of efficiency, indicating the applicability of our solution to real time problems.

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

1. Introduction

Natural color scenes, such as outdoor 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 [1,2,3] 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 pa-

rameters 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** in order to present more stable results. This usually demands complex computations to detect segment-correlation clues, such as the usage of 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 for the first time a ground truth-based objective qualitative validation of our previously presented gradient network approach for image segmentation [4] and analyze its performance in conjunction with another method [1], providing a novel color scene segmentation approach that is extremely fast. It is intended to be used as a two-step approach that shows satisfactory results when applied to natural color scenes, while not showing poorer performance than state-of-the-art methods. Both empirical validations, the objective qualitative analysis as well as the performance comparison are performed against state-of-the-art segmentation algorithms an presented in detail.

The approach described here is based on a pipeline of two fast segmentation algorithms:

- I. A hierarchical region-growing segmentation [1] that generates an over-segmented picture where natural boundaries of objects are preserved.

- II. A color gradient based region-growing post-segmentation method [4] that starts from an initial pre-segmented image and computes a gradient network that spans the entire image. It finds locally connected gradients that show an organized pattern, representing ordered color variations of a same object. These "organized segment clusters" composed of correlated shades of color and light are merged into meta-regions that are then presented as the final segmentation result. One of the most important features of this algorithm is its computational efficiency.

In this paper we present a short review of the *Gradient Network Method* (GNM), highlighting the characteristics responsible for its performance. The *Color Structure Code* (CSC) method is also addressed, followed by results obtained with this combined approach. A processing time comparison is presented taking into account two different combinations of pre-segmentation region growing algorithms with the GNM. It should clarify the main reason

for choosing the CSC as the first stage of this approach. These results are compared against a set of state-of the-art region-growing segmentation algorithms and the results are presented.

Finally we present a discussion that takes into account the applicability of such an algorithm for real-time color image segmentation.

2. Objectives and Validation Methodology

The GNM produced promising results from the qualitative point of view. Since it was developed to provide a stable solution for the segmentation of images containing objects presenting changes in shades of color, while maintaining good performance characteristics, we devised a set of experiments with the objective to demonstrate empirically that the GNM in conjunction with a fast pre-segmentation method is capable of providing an alternative to state-of-the-art segmentation algorithms that is much faster while showing at least the same robustness and segmentation quality levels.

We employed the following method:

- a) Selected a set of state-of-the-art segmentation algorithms to compare to the performance and segmentation results achieved with GNM. This set is different from the one used in [4].

- b) Followed a new empirical ground truth-based validation strategy intended to provide concrete objective qualitative results.

For this purpose, we devised the following procedure:

- I. Two well-known segmentation methods to be used as pre-segmentation procedures were chosen: CSC [1] and Mumford-Shah [3].

- II. Segmentations performed with each of these algorithms were compared against ground-truths using Rand [5] and Bipartite Graph Matching (BGM) [6] indexes. For each segmentation method we selected a wide range of segmentation parameters, selected the result considered to be the best one for every pair of image set and segmentation algorithm and generated Rand and BGM scores for the complete set of ground-truths for each image

- III. These two segmentation methods were also selected to generate over-segmented images to be used in combination with both the original GNM and the version presented in this paper. For each algorithm we selected a set of segmentation parameters that produced over-segmented images where no segment leakage with respect

Aldo von Wangenheim,
et alli

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

to any ground truth was allowed. Each of these results was used as an input for the GNM algorithm, which also was run with a set of different parameters. The resulting segmentations after post-processing with the GNM were also compared against ground-truth images using Rand and BGM indexes.

IV.   We compared these results to two other well-established segmentation methods: JSEG [2] and EDISON [7] also using the ground-truth images and the Rand and BGM indexes.

In this context, the comparison against the EDISON method is new and was chosen because of its quality and robustness. The subjective comparisons against the RHSEG [8] method shown in [4] where not included into this experiment because RHSEG needs user interaction, which can introduce a subjective bias, and is therefore not suited to be tested with the proposed validation methodology.

For the pre-segmentation step, the parameter used for CSC was threshold = 30 and for Mumford-Shah images were generated with lambda = 600. The parameter ranges and increment steps used for these segmentation methods were the following:

1.   CSC: 20 $\leq$ threshold $\leq$ 100, threshold-step = 10;

2.   Mumford-Shah: 1000 $\leq$ lambda $\leq$ 15000, lambda-step = 500;

3.   EDISON: 3 $\leq$ SS $\leq$ 30, SS-step = 1, SR = 8.

4.   JSEG is an unsupervised technique and does not require parameters.

The chosen segmentation quality measures for qualitative validation [5,6] are described in the Appendix. Each measure was chosen because it is a representative of one of two widely employed segmentation quality metrics: *counting of pairs* and *set matching*.

3. GRADIENT NETWORK METHOD

The GNM [4] was developed to deal with segmentation problems where objects in the scene are represented by gradually varying color shades, as they often are found in outdoor scenes. This technique employs a novel segmentation strategy and was developed for robust and fast image segmentation.

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 relat-

ed to the approach in [9]. 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 were performed combining it with a Mumford-Shah functional (MS) based pre-segmentation provided by the Megawave package [10] and with the Color Structure Code (CSC) [1] algorithm. Other techniques, such as Watershed [11], could also be employed for the pre-segmentation step. The pre-segmentation algorithm must only fulfill the requirement of producing super-segmented results that preserve the main edges. The quality of the pre-segmentation, however, affects the final result as is shown below.

The external pre-segmentation step is followed in GNM by a labeling procedure to convert the segmentation output into a graph $G(V, E)$. The next step is to check 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* [12] variations. This allows a better evaluation of the similarity in presence of different luminance in the regions. Regions of continuous and smooth gradients are due to the presence of lighting effects in the scene of an image. With this additional feature, the algorithm becomes more robust when applied to images with 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 [4].

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 high-level structural description of the algorithm can be found in Figure 1.

3.1. BRIEF ALGORITHMIC DESCRIPTION OF THE GRADIENT NETWORK METHOD

A summarized version of the GNM algorithm is given below:

1.   Given a segmented image, a labeling process will be applied and the homogeneous objects in the image will identified.

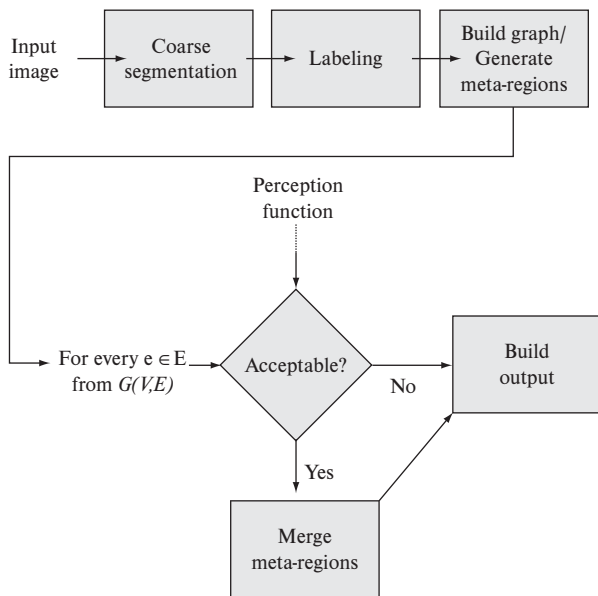2.   The labeled objects have their neighborhood identified. Build a connected graph $G(V, E)$.

Aldo von Wangenheim,
et alli

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

Figure 1. Diagram displaying the GNM process.

3. Associate each vertex $v \in V$ with an unique *meta-region m*, that will be used to represent and group similar regions and that have a path connecting them.

4. With the connected graph, select any edge $e \in E$ of the graph $G(V, E)$.

   4.1. For the current edge, take the two vertices $v_1, v_2 \in V$, and verify if they are not already contained in the same meta-region m. If so, proceed to 4.4. Else, continue.

   4.2. Identify which type of perception, clear or rough, applies to the gradient of color between vertices $v_1, v_2 \in V$.

   4.3. According to the identified perception, evaluate through a similarity measure if the gradient between these vertices is smaller than the threshold defined for the current perception. If it is, the meta-regions m_1 and m_2 containing each of the vertices are merged into a new meta-region m_n. Otherwise, do nothing.

   4.4. Mark this edge as *verified*. Select an edge that has not been yet verified and go back to the first step of this group of instructions. If there is none, follow to step 5.

5. With the meta-regions found in the former steps build the output image, as each meta-region now represents what is considered an object in the scene of the image by the algorithm. Represent

each pixel in the a meta-region by the mean value of the pixels of the meta-region.

For a more formal description of this algorithm we refer to [4].

4. THE COLOR STRUCTURE CODE

The Color Structure Code (CSC) [1] 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 real-time road sign recognition. The CSC is a region growing algorithm that uses a hierarchical topology formed by islands, a topology type introduced by [13]. These islands have different levels, as shown in Figure 2. A level 0 island is a hexagon, composed by the 6 vertex points around a central point. During the process, some islands overlap others such that level $n + 1$ islands are composed by seven level n overlapped islands. This will be repeated until an island spans the entire image.
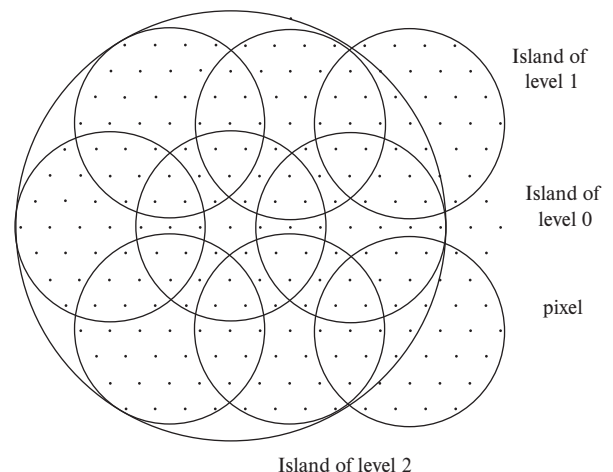

Figure 2. CSC's hierarchical island structure. [1]

As a first step, the whole image will be partitioned into level 0 islands. A merging step, where the islands will grow and overlap iteratively, will follow. After the grouping step, a split step is performed, 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.

The CSC approach alone is a suitable alternative for dealing with outdoor scene images. However, it is not prepared to deal with regions that present larger continuous color or luminance gradients, where segmentation problems will occur, as shown in Figure 3. As a usual conse-

Figure 3. CSC results when processing a complex outdoor image. From left to right: the original image, a cautious approach and an aggressive approach. See color plate in page 40.

quence, found in most algorithms, sensitive regions might be swallowed or more cautious parameters might produce many more segments than would be useful. We will show that the CSC shows good performance and reliability regarding outdoor images when employed as a pre-processing step for the GNM, which then performs the sensitive global gradient-based region grouping actions.

5. ACHIEVING EFFICIENCY AND ROBUSTNESS

High speed performance segmentation algorithms have been investigated to satisfy the demands from applications that require real-time results. Fast segmentation processes could be used in several situations, like motion detection in video frames or autonomous vehicle guidance [1]. Another application would be to guide surgery and other medical procedures. An example is given in [14], where segmenting a carotid artery is a useful step in medical imaging. Efficiency requirements can also be found in several other areas. [15] presents a technique developed for real time applications as space-weather analysis. In [16] an approach is proposed to track players in a soccer pitch. Fast segmentation approaches are a recurrent topic and several optimizations or specializations over known techniques have been developed [17,18,19].

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

5.1. 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 are done together with the graph generation step.

To merge regions presenting similar *perception* [12], the graph is then traversed Since this step depends solely on the 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.

5.2. PRE-SEGMENTATION ISSUES

As our main focus here is to obtain robust results combined with high performance, we have chosen the *Color Structure Code* (CSC) [1] 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 general robustness 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 quality and performance comparison, we used the traditional Mumford-Shah Functional implementation supplied by the *Megawave* [10] image processing package. The behavior of this method is well known and documented and was considered for a long time the best choice for quality comparisons.

6. RESULTS AND DISCUSSION

To empirically validate the approach, 17 outdoor images showing different color and texture characteristics where processed with all methods. To allow performance comparisons, all tests were run on the same computer.

The adoption of the Berkeley's image dataset [20] 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 combined segmentation techniques used in the following tests were: GNM applied over pre-segmented images by CSC, with a threshold equal to 30 and GNM applied over pre-segmented images by a Mumford-Shah

functional based segmentation, with lambda equal to 600. The GNM parameters were iterated over a range of reasonable values and the results showing the best Rand and BGM indexes were chosen.

For all other four results, CSC and MS alone, and JSEG and EDISON, the parameters were not preset, but were iterated over a range of reasonable values and the results showing respectively the best Rand and BGM indexes were chosen. It was allowed for the best results according to Rand and BGM to be different.

6.1. PERFORMANCE

The mean execution time for all images with each method is shown in Figure 4. The total execution time for each set with every selected algorithm is shown in Table 1. 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.
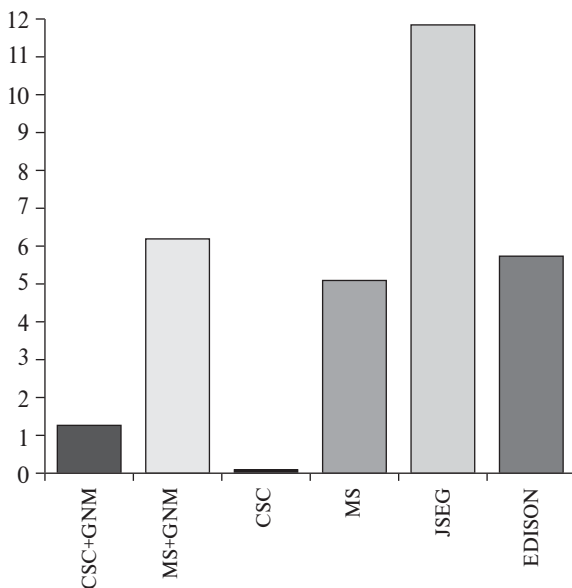
Figure 4. Chart comparing mean execution times in seconds.

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. Figure 5 shows image results obtained with GNM combined with both CSC and Mumford-Shah. As the Table 1 shows, the combination of CSC and GNM shows results with a mean value of about 1.2 seconds, which is only slower than CSC. This was expected, considering the cumulative times of CSC+GNM.

The mean time for CSC+GNM is several times shorter than Mumford-Shah (including MS+GNM), EDISON and JSEG, which is the slowest of all. There is little standard deviation among the times obtained for CSC+GNM, while

again in accordance with the exception of CSC alone, all other techniques show higher standard 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 several small image 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, that GNM has a very stable performance in both cases, with little deviation among the tests cases.

6.2. QUALITY

Figure 5 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 http://www.lapix.ufsc.br/fast .

Tables 2 and 3 show the results of the objective qualitative validation of the segmentations using the RAND and BGM indexes. We have included the mean results for each image and method for comparison and the standard deviation as a measure of robustness.

For the RAND index both combinations of the GNM with pre-segmentations showed the best mean results and also the highest robustness, presenting the lowest standard deviation among the results.

CSC+GNM is the second best, being behind only MS+GNM and slightly better than EDISON.

For the BGM index, the combinations of the GNM with pre-segmentations scored at place 3 and 4. The best results were achieved by the EDISON method, being followed closely by the CSC alone.
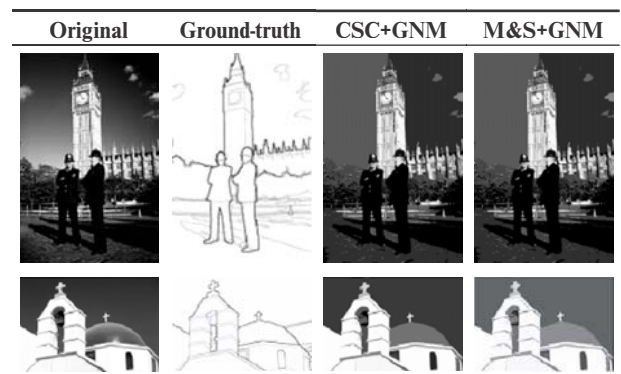

Figure 5. Examples of the image results obtained by GNM, combined with both CSC and Mumford-Shah. The first row corresponds to image number 368078 and the second row corresponds to 118035 from Berkeley image dataset. See color plate in page 40.

Table 1. The total execution times for 17 Berkeley dataset images processed using the techniques is listed in Results and Discussion. In the GNM cases, the time consumed by GNM is shown in parenthesis. Landscape images dimensions are 481 x 321 pixels and portrait images are 321 x 481 pixels. The mean time, standard deviation in seconds and standard deviation in percent for every algorithm test are displayed in the last lines of the table.

Image	Execution Times in Seconds CSC + GNM	MS + GNM	CSC	MS	JSEG	EDISON
2092	0.625 (0.547)	5.422 (0.500)	0.078	4.782	9.30	4.52
3096	0.719 (0.641)	5.313 (0.547)	0.063	4.641	7.49	2.15
14037	0.625 (0.547)	9.203 (0.516)	0.062	8.172	11.14	4.81
15088	0.656 (0.547)	5.547 (0.515)	0.078	5.063	26.50	7.14
22090	0.609 (0.515)	5.187 (0.531)	0.078	4.687	9.06	5.12
24004	0.719 (0.625)	5.609 (0.609)	0.094	4.969	10.78	6.00
24063	0.594 (0.516)	5.203 (0.453)	0.062	4.828	9.08	3.07
46076	0.641 (0.547)	5.329 (0.516)	0.078	4.766	11.09	5.89
48055	0.828 (0.719)	5.563 (0.641)	0.094	4.828	14.22	7.18
67079	0.672 (0.578)	5.360 (0.516)	0.079	4.937	12.33	6.77
118035	0.702 (0.593)	6.093 (0.531)	0.078	5.375	9.39	3.78
124084	0.657 (0.563)	5.109 (0.562)	0.078	5.375	9.39	11.39
138078	0.750 (0.656)	6.030 (0.562)	0.078	5.360	14.36	5.53
143090	0.641 (0.547)	5.047 (0.578)	0.078	4.469	12.91	4.41
219090	0.719 (0.625)	5.359 (0.593)	0.093	4.735	11.84	5.23
310007	0.641 (0.531)	5.609 (0.500)	0.094	5.031	9.89	7.28
368078	0.734 (0.672)	5.093 (0.593)	0.078	4.547	12.58	7.20
mean	0.678 (0.586)	5.652 (0.545)	0.079	5.092	11.844	5.734
std.dev.	0.061 (0.06)	0.962 (0.047)	0.010	0.840	4.240	2.096
% std.dev.	9% (10%)	17% (9%)	13%	16%	36%	37%

Table 2. Segmentation quality results for 17 different Berkeley Dataset outdoor images according to the Rand Index

Image	RAND Index CSC + GNM	MS + GNM	CSC	MS	JSEG	EDISON
2092	0.056	0.106	0.091	0.231	0.152	0.124
3096	0.265	0.155	0.150	0.172	0.347	0.132
14037	0.168	0.134	0.267	0.112	0.128	0.145
15088	0.119	0.119	0.133	0.278	0.481	0.120
22090	0.049	0.051	0.051	0.141	0.054	0.108
24004	0.183	0.184	0.188	0.237	0.232	0.200
24063	0.176	0.159	0.147	0.238	0.270	0.146
46076	0.105	0.123	0.194	0.112	0.079	0.096
48055	0.071	0.081	0.121	0.151	0.153	0.078
67079	0.294	0.283	0.220	0.303	0.242	0.283
118035	0.109	0.102	0.108	0.096	0.107	0.105
124084	0.281	0.304	0.293	0.312	0.284	0.270
138078	0.184	0.189	0.109	0.157	0.118	0.052
143090	0.091	0.090	0.316	0.218	0.235	0.094
219090	0.154	0.060	0.167	0.192	0.083	0.048
310007	0.097	0.112	0.429	0.134	0.227	0.437
368078	0.079	0.079	0.235	0.127	0.138	0.073
mean	0.146	0.137	0.189	0.189	0.196	0.148
std.dev.	0.077	0.071	0.096	0.069	0.110	0.100

Table 3. Segmentation quality results for 17 different Berkeley Dataset outdoor images according to the BGM Index.

Image	BGM Index					
	CSC + GNM	MS + GNM	CSC	MS	JSEG	EDISON
2092	0.124	0.129	0.114	0.262	0.056	0.103
3096	0.138	0.033	0.010	0.045	0.220	0.048
14037	0.175	0.340	0.017	0.187	0.275	0.026
15088	0.073	0.109	0.084	0.170	0.375	0.114
22090	0.243	0.284	0.242	0.147	0.214	0.204
24004	0.430	0.463	0.412	0.557	0.555	0.424
24063	0.104	0.089	0.037	0.306	0.326	0.054
46076	0.318	0.409	0.245	0.331	0.259	0.072
48055	0.114	0.139	0.198	0.336	0.402	0.081
67079	0.104	0.124	0.217	0.145	0.213	0.117
118035	0.202	0.275	0.192	0.205	0.238	0.066
124084	0.610	0.702	0.279	0.329	0.624	0.388
138078	0.394	0.419	0.147	0.298	0.244	0.053
143090	0.206	0.206	0.173	0.465	0.512	0.157
219090	0.338	0.211	0.205	0.314	0.214	0.195
310007	0.137	0.181	0.086	0.292	0.340	0.038
368078	0.355	0.385	0.154	0.489	0.455	0.314
mean	0.239	0.264	0.165	0.287	0.325	0.144
std.dev.	0.147	0.172	0.103	0.133	0.147	0.123

7. Conclusions and Discussion

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 methods that were available for comparison when this paper was being written. On the other side, the 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 Method step was used in combination with more traditional segmentation approaches such as the Mumford-Shah functional.

Considering that CSC + GNM is five times faster than EDISON, which is the method that has shown the best quality scores as a standalone approach, it is noticeable that the CSC + GNM segmentation quality scores are so high, being even better than the EDISON according to the Rand index. CSC + GNM also presented a stable behavior, both in performance, showing little variation in processing time, and also in quality, showing little variation in the Rand index, thus providing extremely robust image segmentation results.

The EDISON implementation provided the best BGM index scores, while JSEG provided the worst ones. Both combined GNM approaches remained in the middle. This is a good enough score, when considering that the GNM is being compared to state-of-the-art segmentation methods. On the other hand, as the results provided in http://www.lapix.ufsc.br/fast show, the BGM method tends to prefer bigger regions, even if a region overlaps partially into another. This leads to results showing under-segmentations with segment leakage receiving a higher score than with the Rand method. We have developed our method explicitly to analyze rigorously region borders in order to avoid such leakages, even if some over-segmentation is left behind. This can be one reason for the poorer performance according to the BGM index. This would also explain why the CSC alone scored second best and even showed the best robustness according to this same validation index.

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 interesting to note that the quality of the final results is 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 process-

Aldo von Wangenheim,
et alli

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

ing 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, even if the pre-processing method is not so good as more traditional approaches.

From the performance point of view, we did not analyze the methods under varying parameter settings, even if we processed each image under approximately 30 different parameter settings for each method except JSEG. Our focus was quality with speed, thus we considered only the segmentation result which presented the best quality under each of the metrics to compute the performance. So, the processing time shown in our result tables is always the time it took to process the segmentation that showed the best quality score.

Further improvements, however, 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.

When the experiments described in this paper were being performed, a new variant of the Mumford-Shah algorithm was published, that is described as overcoming one of the most important shortcomings of Mumford-Shah, namely the long processing time [21]. We did not have the opportunity to implement and test this new variation, but since CSC is in average 64 times faster than the traditional Mumford-Shah, showing a mean segmentation time of 0.079s compared to the mean segmentation time of 5.092s presented by the standard Mumford-Shah implementation, while presenting the same mean Rand index of 0.189 and a better BGM index of 0.165 against the BGM index of 0.287 of Mumford-Shah, we think it is still a better choice for the pre-segmentation step, even if faster versions of the Mumford-Shah algorithm are appearing.

REFERENCES

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

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

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

[4] A. V. Wangenheim, R. Bertoldi, D. Abdala, M. M. Richter. Color image segmentation guided by a color gradient network. *Pattern Recognition Letters*. 28: 1795-1803, 2007.

[5] W. M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of American Statistical Association*. 66: 846-850, 1971.

[6] X. Jiang, C. Marti, C. Irniger, H. Bunke. Distance measures for image segmentation evaluation. *EURASIP Journal on Applied Signal Processing*. 1-10, 2006.

[7] D. Comaniciu, P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24(5): 603-619, 2002.

[8] J. C. Tilton. D-dimensional formulation and implementation of recursive hierarchical segmentation. Disclosure of Invention and New Technology: NASA Case N'I GSC 15199-1, May 2006.

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

[10] Megawave image processing package. http://www.cmla.ens-cachan.fr/Cmla/Megawave/, Sept 2006.

[11] L. Vincent, P. Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 13: 583-598, 1991.

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

Aldo von Wangenheim,
et alli

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

where n is the cardinality of P and n_i and n_j are the cardinality of the clusters $c1_i$ and $c2_j$.

A.2   BIPARTITE GRAPH MATCHING

The *BGM index* [6] computes an one-to-one correlation between clusters at the same time trying to maximize their relationship. It considers each cluster of the C1 and C2 clustering as vertices of a bipartite graph. Edges are added between each vertex of the two partitions and they are valued as $|c1_i \cap c2_j|$, a value that can be directly extract from the matching matrix. Then the maximum-weight bipartite graph is defined as the subgraph $\{(c1_{i1}, c2_{j1}), ..., (c1_{ir}, c2_{jr})\}$ where only the edges from $c1_i$ to $c2_j$ with maximum weight are present. After all max-valued edges were found the overall graph weight is calculated by sum of all remaining edge weights.

$$BGM(C1, C2) = 1 - \frac{w}{n} \qquad\qquad A.IV$$

COLOR PLATE

Figure 3. CSC results when processing a complex outdoor image. From left to right: the original image, a cautious approach and an aggressive approach.
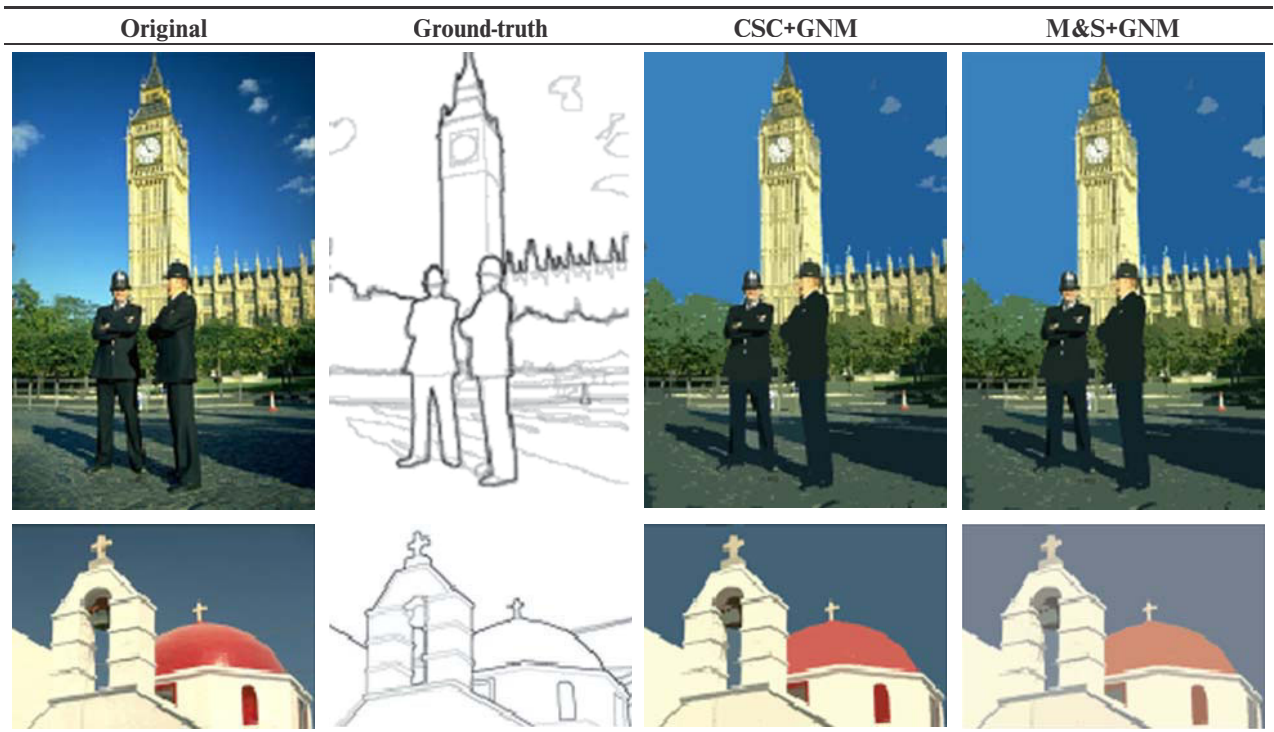

Figure 5. Examples of the image results obtained by GNM, combined with both CSC and Mumford-Shah. The first row corresponds to image number 368078 and the second row corresponds to 118035 from Berkeley image dataset.