Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 1795–1803
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Color image segmentation guided by a color gradient network

Aldo v. Wangenheim [a,*], Rafael F. Bertoldi [a], Daniel D. Abdala [b], Michael M. Richter [c]

[a] Image Processing and Computer Graphics Lab – LAPIX, CS Department, Federal University of Santa Catarina – UFSC, Florianópolis, SC, Brazil
[b] Computer Vision and Pattern Recognition Group, University Münster, Münster, Germany
[c] CS Department, University of Calgary, Calgary, Alberta, Canada

Received 20 December 2006
Available online 26 May 2007

Communicated by H.H.S. Ip

Abstract

Existing region-growing segmentation algorithms are mainly based on a static similarity concept, where only homogeneity of pixels or textures within a region plays a role. Typical natural scenes, however, show strong continuous variations of color, presenting a different, dynamic order that is not captured by existing algorithms which will segment a sky with different intensities and hues of blues or an irregularly illuminated surface as a set of different regions. We present and validate empirically a new, extremely simple approach that shows very satisfying results when applied on such scenes, while not showing poorer performance than traditional methods when applied to standard region-growing problems.

Keywords: Region-growing segmentation; Natural color scenes; Color gradient networks

1. Introduction

Color image segmentation algorithms present a very important application field of both computational image analysis and computer vision. Their main goal is to identify homogeneous regions or regions of similar characteristics within an image according to a predefined aspect, especially with respect to their color intensities. This family of algorithms becomes very important when we intend to achieve a result that one could compare closely to the human vision.

An essential problem of current algorithms is that similarity identification is calculated over simple continuous pixel neighborhood similarity. This approach leads to errors when we consider, for instance, the blue sky on a cloudless day. The sky presents a darker blue intensity on the upper parts and gradually and continuously decreases into lighter blues on the lower parts. This intensity variation will be identified as numerous regions on most of the current algorithms, while the human eye sees only one region, the sky itself, recognizing the order in the variation. The concept of an image segmentation algorithm based on the analysis of the relationships of continuous gradients aims to overcome this difficulty. We call this approach the gradient network segmentation method and this paper presents an overview of the method and results so far.

1.1. Color gradient network-based segmentation

Traditional region-growing segmentation algorithms, in order to produce the best possible segmentation of an image, *i.e.*, one that best divides the image, rely on the search for islands of high entropy in the image, *i.e.*, homogeneous regions where no changes occur. The Mumford and Shah segmentation algorithm (Mumford and Shah, 1989) is the best example of this methodology. The more

* Corresponding author. Present address: LAPIX/LISHA, Departamento de Informática, CTC, UFSC, Campus Universitário, 88.049-900 Florianópolis, SC, Brazil. Tel.: +55 48 3331 9516; fax: +55 48 3331 9516 x16.

E-mail address: awangenh@inf.ufsc.br (A.v. Wangenheim).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.05.009

common Watershed algorithm Vincent and Soille (1991) and its more elaborate implementations, such as (Mittelheusser and Kruggel, 1995), although they use image gradients to define region boundaries, are not much different. This leads to a situation where images that show a clear order, but where this order is represented by the change patterns of colors or light intensities, are poorly segmented, see Fig. 1.

On the other side, when we humans observe a scene with a blue sky, where the blue tones change gradually from the horizon on, the order is obvious and we do not even try to interpret the sky as a series of regions of different "blueness". *We recognize the order in the dynamics of the colors* and identify the sky as one single structure in the image. This occurs also with similar images where we do not have background knowledge, such as colored curved shapes that have never been seen before, where an irregular illumination produces different shades of color and light intensity. But traditional segmentation algorithms are not able to perform this simple and foreknowledge-independent recognition operation. This is illustrated by the scene (b) in Fig. 1, segmented with the color space implementation of the Mumford and Shah algorithm found in the new release of the Megawave package (Megawave, 2006), where the result shows the lake over segmented. If we modify the parameters in such a radical way that the sky or the lake is segmented as one single region, this will produce segment leakages in other parts of the image, resulting in an unusable segmentation. If we, however, take the gradual change patterns in the image into account, we will be able to obtain a segmentation result as shown in Fig. 1c.

Several approaches have been developed in the last 15 years to enhance the performance of traditional region growing segmentation techniques on natural scenes, ranging from simple hierarchical approaches (Priese and Rehrmann, 1993; Rehrmann and Priese, 1998) to very elaborate ones (Dupuis and Vasseur, 2006). More recent methods focus not only on identifying very similar color regions, but also on additional features such as the texture of objects (Deng and Manjunath, 2001; Dupuis and Vasseur, 2006; Kato and Pong, 2006). But even with the use of techniques capable of identifying objects of similar color and even texturized objects, there is another characteristic that troubles segmentation algorithms: luminosity and the existence of large but slow and gradual variations of the color of the same object in an image, resulting in high total gradients of low local steepness. Some approaches have tried to solve the problem using different evaluation methods for chromacity and luminosity (Dony and Wesolkowski, 1999; Schneider et al., 2000).

We propose a novel post-segmentation technique to process pre-segmented images through the analysis of gradient networks that describe color variation patterns between segments, a method especially robust when applied to large continuous color variations in the same object in an image, as occurs with typical outdoors images. The objective of this method is to perform reliable and accurate segmentations of objects that show strong continuous color variations, as are caused by varying light incidence on outdoors color images.

2. Methods

2.1. General principle

The goal of our approach is to describe the dynamics of color-changing and to use them in a segmentation algorithm, trying to *recognize a higher order of organization in the image*, namely the *organization of the changes* or the *structure of the tendencies* in the image. The general idea is to perform this task using a graph that represents the nature of modifications between regions of the image as a piecewise cost function. This graph is then processed through an algorithm that melts regions that pertain to the same image-dynamic process, *i.e.*, that lay within a tendency, such as a long upward or downward slope. The algorithm was developed to be able to represent the chromatic differences between regions as costs and to recognize order in the change patterns of the image, melting regions that pertain to the same "ordered structure" of the graph. The method we propose is a general framework, and some of the components cited below could be changed through other techniques.

Initial image representation: To represent each single pixel as a node in a graph with a set of parameters is a costly solution and may introduce local noise that does

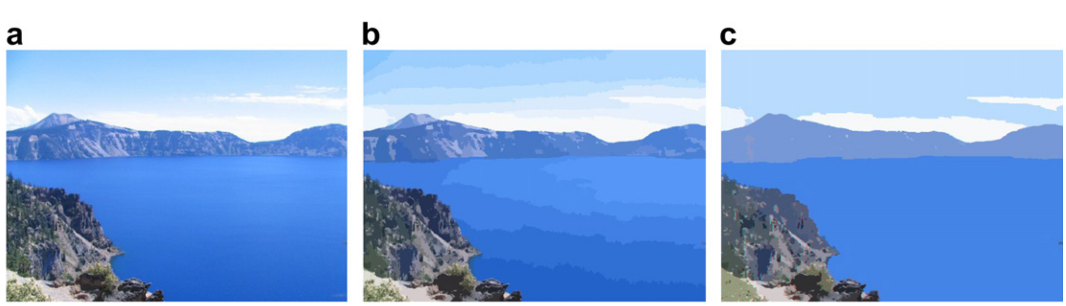


Fig. 1. A well-organized image where the order is given by structured gradients rather than by abrupt changes and two different segmentations of the image: (a) original image, (b) Mumford and Shah segmentation of the image using (Megawave, 2006), (c) Gradient network segmentation of the same image.

not contribute to the regional estimation of the color variation tendency. Thus we use region-segmentation as a pre-processing step, which will represent each region that is statically, *i.e.*, value-wise homogeneous within a set of parameters as a single object. This set of parameters will be domain- and pre-segmentation method-dependent and is discussed below.

Each object generated by this pre-segmentation will be considered an *atomic region*. This segmentation should be performed with such parsimonious parameters that the resulting image is over segmented and no segment leakage occurs. The gradient network is then built upon this over segmented image, with each *atomic region* being represented as a node in the gradient graph. In our implementation we have chosen the Mumford and Shah functional (Mumford and Shah, 1989) for the pre-segmentation because of its robustness and quality of results, but one could use a faster method such as Split and Merge, Watershed (Vincent and Soille, 1991) or even the CSC (Priese and Rehrmann, 1993), since one does not search for a final segmentation, but for a pre-classification of pixels in homogeneous structures. For our validation, we tried to devise a set of stable parameters for the Mumford and Shah algorithm that did not produce any segment leakages and to use them throughout the experiment.

Similarity measure: What is regarded as homogeneous will be defined by a threshold. In principle, the similarity measure compares two such regions considered currently as homogeneous, measuring how close the regions are with respect to the degree of homogeneity. These regions are represented in a graph and the role of a similarity measure is to provide a cost for each edge connecting adjacent regions. The details will be presented in Section 2.2.

As in the case of the initial segmentation, the similarity measure calculation is also a slot in the framework where many different similarity measures could be used. In the experiment described in this paper, we chose to differentiate regions of clear and rough color perception in the scene of the image, as originally discussed by (Huang et al., 2006), improving the robustness of the color similarity measure to specular lights and shadows. The idea was to enhance robustness in the presence of strong luminosity variations, something usual with real world images. Besides the similarity measures, the thresholds will play an important role. A more formal description of this algorithm will be given below.

2.2. Description of the gradient network method

(1) A pre-segmented image is used as input data. It is produced with such conservative and oversensitive parameters as to avoid that a segment in this image leaks from one object into another in the image.
   (i) Along with the input image, two thresholds are passed as parameters. These thresholds, t_{cp} and t_{rp}, will be used to verify if the result of a gradient of colors applied to the similarity function are acceptably smooth according to the perception associated to the threshold used.
   (ii) All segments in this input image are labeled.

(2) Regions and their neighborhood relationships are represented through a connected graph $G(V,E)$. The labeled regions from the pre-segmented input image will be the vertices V of the graph and the neighborhood relations among those objects will be represented as the edges E of the graph. V will be called the *Set of Atomic Regions*.
   (i) This graph will always be a connected graph because all regions represented by the vertices are contained in the same image, and through the edges that link neighbor regions it is possible to reach from any given vertex any of the other vertices as it is possible to find a path from any pixel p_1 to any other pixel p_2 in a same image.
   (ii) Then, given the graph $G(V,E)$ of connected regions, it can be defined as follows:

$$V = \{v \in V | \forall((x_1,y_1) \subset v) \wedge (\neg\exists((x_2,y_2) \subset v) \wedge \mu(x_1,y_1) \neq \mu(x_2,y_2))\}$$

   where every vertex $v \in V$represents an Atomic Region and contains a set of ordered pairs of coordinates (x,y) in the orthogonal axis X and Y. For all these contained ordered pairs of coordinates (x,y), the positions they all represent in the labeled input image will all have a same and unique value, where μ is the function that gives the label value for a given pair of coordinates (x,y) in the labeled image.

$$E = \{e \in E | (v_1, v_2 \in V) \wedge (v_1 \neq v_2) \wedge \exists(((x_1,y_1) \subset v_1) \wedge ((x_2,y_2) \subset v_2) \wedge (N_8(x_1,y_1) \supset (x_2,y_2)))\}$$

   where every edge $e \in E$ is formed by a pair of vertices $v_1, v_2 \in V$ and between these two vertices there is at least pair of positions $(x_1,y_1) \subset v_1$ and $(x_2,y_2) \subset v_2$ that show an 8-connected kind of neighborhood.

(3) In the next stage of the algorithm, each vertex $v \in V$ is associated with a new and unique *Meta-region* $m \in M$. Meta-regions are logical containers that will be used to store groups of regions that have a path among them and that are linked through color gradients that are considered acceptable according to the similarity measure proposed. The goal of the gradient network method is to merge meta-regions.
   (i) All meta-regions will behave like connected sub graphs, meaning that any vertex contained by a given meta-region has at least a path to any other vertex also contained in the same meta-region.

$$M = \{m \in M | v_1, v_2 \in V \wedge v_1, v_2 \subset M \wedge f_{\mathrm{c}}(v_1, v_2) \leqslant T_{\mathrm{c}}\}$$

   where every meta-region $m \in M$ is formed by a set of vertices $v \in V$ that have an acceptably smooth gradient between pairs of vertices $v_1, v_2 \in V$ according to the T_{c} threshold.

(ii) Any vertex v_i can pertain to only one Meta-region.

(4) The algorithm then runs through the graph $G(V,E)$ looking in the edges $e \in E$ that connect vertices $v_1, v_2 \in V$ for gradients smooth enough to be considered acceptable according to the perception for this edge, resulting in the union of those vertices in the same meta-region $m \in M$.

(i) The goal is verifying if these two regions can be placed in the same meta-region m, meaning that both are considered similar by the measure. If the vertices already are contained by the same meta-region $m \in M$ there is no reason to keep verifying and therefore the process goes on, proceeding to another unverified edge, if there is still one. The edge is marked as *verified*. Each edge is examined only once.

(ii) If both vertices are not in the same meta-region $m \in M$, then the kind of perception to apply to this link will be determined.

The perception idea was developed based on concepts found in (Huang et al., 2006), where a difference of color perception is used to enhance the contrast evaluation in images. With these perceptions and the features they provide, it is easier to make the method robust to change of lighting and presence of shadows in the scene. Also it is an interesting concept because the question of smooth variation of colors in objects is directly connected to illumination and this type of evaluation through perceptions offers a powerful tool for dealing with this type of characteristic.

In our algorithm we use two kinds of perceptions:

(1) *Clear perception*, where there is good color saturation and average levels of luminosity.

(2) *Rough perception*, where there is low color saturation, or very high or very low intensities.

The color values of the vertices are converted to the HSL color model, composed by three orthogonal axes that represent, respectively, hue, saturation and luminance. Currently the acceptance levels for both perceptions are defined parametrically as ranges of acceptance for the hue, saturation and luminance values, accordingly to (Huang et al., 2006). These ranges must partition the range that HSL values can assume in a way that a region is either classified as having a clear or rough perception. These parameters will allow a better configuration of the segmentation for the lighting conditions of the scene of the image.

The equations need two vertices as input arguments. These will always be neighboring regions, as a consequence from the fact that the input vertices are extracted from an edge of the graph and inherits this property from this edge. Given both vertices and the kind of perception their gradient fit, it is checked if the value found with the evaluation function f_c is acceptable according the threshold T_c defined parametrically for the perception used. The function definition is

$$f_c(v_1,v_2) = \begin{cases} \alpha_{cp} * \min\left(\left|\int_{H(v_1)}^{H(v_2)} d\theta\right|, 1.0 - \left|\int_{H(v_1)}^{H(v_2)} d\theta\right|\right) \\ \quad +\beta_{cp} * |S(v_1) - S(v_2)| + \gamma_{cp} * |L(v_1) - L(v_2)|, \\ \quad \text{perception}_{\text{clear}} \\ \beta_{rp} * |S(v_1) - S(v_2)| + \gamma_{rp} * |L(v_1) - L(v_2)|, \\ \quad \text{perception}_{\text{rough}} \end{cases}$$

where f_c is the smooth gradients evaluation function and it takes two vertices $v_1, v_2 \in V$ as parameters. The functions H, S, L take a vertex $v \in V$ as parameter and give the mean value of, respectively, hue, saturation and luminance from the color values of the pixel positions in the image that the vertex v contains. The coefficients α, β, γ correspond to the desired relevance given by the perception to hue, saturation and luminance, respectively. These coefficients are also parameters of the algorithm.

(iii) In case of success, meaning that the result is smaller than the threshold T_c, a merging is performed between the meta-regions $m_1, m_2 \in M$ associated with these vertices $v_1, v_2 \in V$, because the gradient between these regions was considered so smooth that these regions actually identify a single object in the scene of the image. No other changes are performed on the vertices or the edge.

$$f_c(v_1,v_2) < T_c \rightarrow m_1, m_2 \in M | v_1 \subset m_1 \wedge v_2 \subset m_2, m_1 \cup m_2$$

If it was not lower than the threshold, then nothing is done and the algorithm continues. Step 4 will be performed once for all the edges present in the connected graph. The resulting meta-regions will represent the objects of the scene, formed by regions of low entropy, found by the initial segmentation that created the input image, that are connected through smooth color gradients.

(5) Finally, the output image will be built based on the meta-regions $m \in M$, where now the color of each resulting Meta-region will correspond to the average color of all vertices $v \in V$ contained in the Meta-region.

3. Results

The gradient network method was applied on reference images provided by (Martin et al., 2001) and a few others collected by the authors. Processing results are shown in Fig. 2. The results were also compared to well-known state-of-the-art region-growing segmentation algorithms,

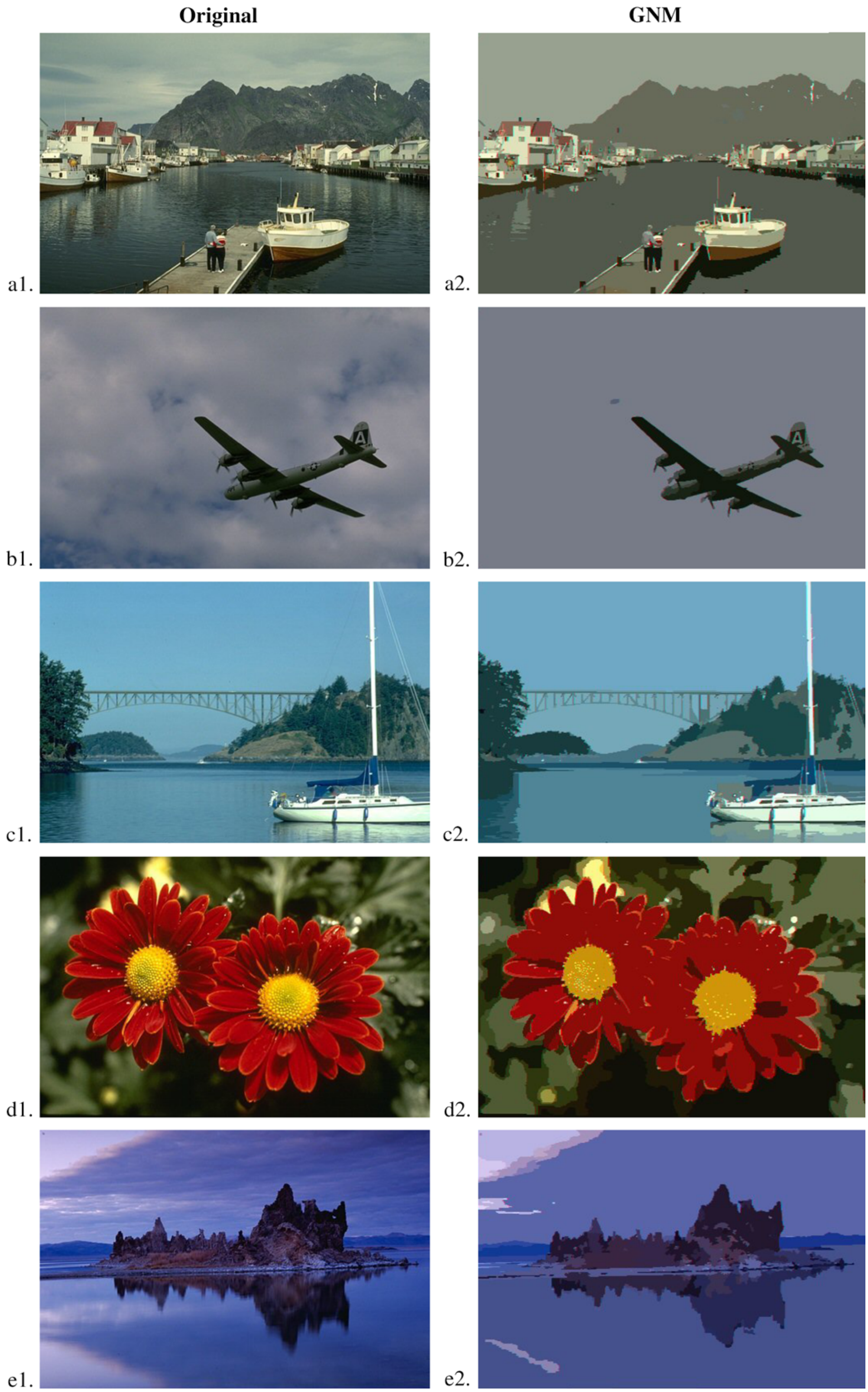


Fig. 2. Results obtained with the gradient network segmentation method on the (Martin et al., 2001) segmentation test images. All tests underwent a Mumford–Shah pre-segmentation with lambda = 600.

namely CSC (Rehrmann and Priese, 1998), Mumford–Shah (Mumford and Shah, 1989), RHSEG (Tilton et al., 2006) and JSEG (Deng and Manjunath, 2001), which were applied to the same images.

The methods used in the comparison shown in Fig. 3 were applied using the following parameters:

- All CSC images were produced with a threshold equal to 50 or 70.
- Each Mumford–Shah was generated with a different and specific lambda value. These lambda values are shown below the images.
- JSEG was used in an unsupervised way with all images.
- All RHSEG images were created with the similarity "entropy" (number 9 in the *params* file – segmentation parameters), with a factor of convergence equal to 1.75 and with a 0.1 importance to spectral clustering.

As stated in the algorithm description, the perception classification parameters must be defined manually. The parameters used for this work were the same for all images, using similar ranges to those found in (Huang et al., 2006). With all HSL values normalized to the [0, 1] range:

$$\text{Perception}_{\text{clear}} = \{(S > 0.1) \wedge (0.2 < L < 0.95)\}$$
$$\text{Perception}_{\text{rough}} = \{(S \leqslant 0.1) \vee (L \leqslant 0.2) \vee (L \geqslant 0.95)\}$$

Depending on the classified perception, the coefficients α, β, γ have the following values, also normalized to the [0, 1] range, for all the tests presented:

$$\alpha_{cp} = 0.7, \quad \beta_{cp} = 0.2, \quad \gamma_{cp} = 0.1$$
$$\beta_{rp} = 0.2, \quad \gamma_{rp} = 0.8$$

A larger set of processing results and comparisons is available at http://www.lapix.ufsc.br/gnm/. All tests shown here can be downloaded from http://www.lapix.ufsc.br/gnm/Images.rar.

The results depicted in Figs. 2 and 3 show that this simple algorithm is capable of organizing a pre-segmented image in meta-regions accordingly to the color variation patterns that a human observer would also see, producing coherent segmentations of objects in images, even when there is considerable variation within a meta-region, given that this variation is relatively smooth and continuous.

4. Discussion

Simplicity was one of the goals kept in mind when this method was conceived. Similar approaches have been used previously by (Trémeau and Colantoni, 2000) and (Wu and

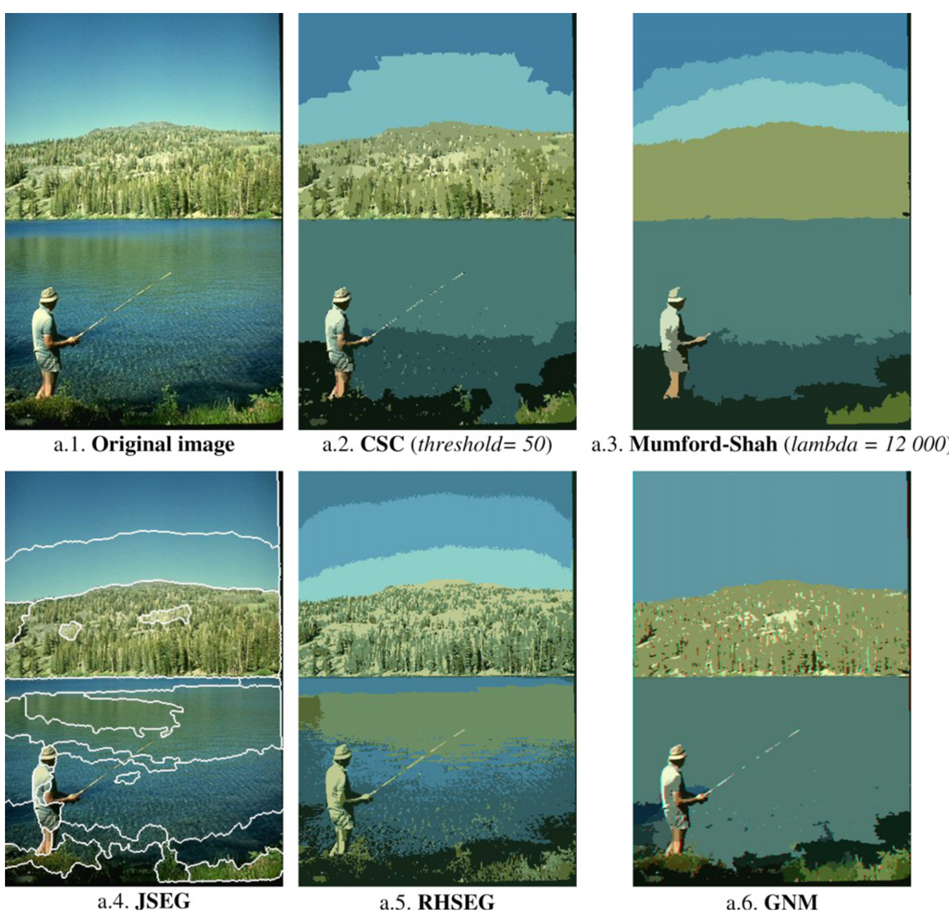


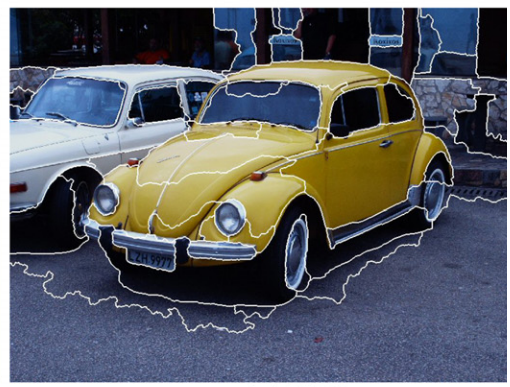
Fig. 3. Comparison of the GNM with other segmentation methods. (a) and (c) are images from the (Martin et al., 2001) segmentation test database. (b) is an image taken by the authors.

b.1. **Original image**

b.2. **CSC** (*threshold = 70*)

Color Image Segmentation guided by a Color Gradient Network

b.3. **Mumford-Shah** (*lambda = 9 000*)



b.4. **JSEG**

b.5. **RHSEG**

b.6. **GNM**

Fig. 3 (*continued*)

Leahy, 1993), where the problem was stated as a graph-flow optimization in sub graphs. Another graph-partitioning approach introduced a global measure obtained by a principal component analysis of a set of different parameters associated with each element of the image (Dupuis and Vasseur, 2006). Our approach also finds sub graphs in the original network but is much simpler and acts only locally on each edge, without trying to find an optimal value for the entire sub graph that represents a region and without using any global image data or cost function. It applies only simple non-adaptative local similarity measures and produces quality segmentation results on difficult images, comparable to the most complex approaches. The method is based on a simple assumption: *local variation patterns of the characteristics of the image are as important as the static characteristics*. To discover order in these patterns is important for a segmentation algorithm. The process of discovering the variation patterns of the image is described as a simple local similarity-based process in one simple and consistent modular framework, with an elegant and integrated definition of stopping criteria.

Our approach can also be seen as a labeling procedure. The goal of labeling techniques is to label uniquely all the objects in an image. Usually it is a process applied after the use of some other image preprocessing method, often a segmentation. However, the result of a region-growing segmentation behaves like a labeling method, assigning pixels to regions and associating mean color values with

c.1. **Original**

c.2. **CSC (*threshold* = *50*)**

c.3. **Mumford-Shah (*lambda* = *12 000*)**

c.4. **JSEG**

c.5. **RHSEG**

c.6. **GNM**

Fig. 3 (*continued*)

them. From this point of view, the pre-processing our approach relies on is a pre-segmentation with a method that points out the obvious region patches where a gradient-pattern analysis is not necessary, thus avoiding unnecessary processing.

The main difference between this method and other, image post-processing labeling methods developed to "correct" over segmentations is that it is not knowledge-based or application-domain-model-based, *i.e.*, it does not need additional knowledge on the content or the application domain of the images, depending only on image information.

References

Deng, Y., Manjunath, B.S., 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Anal. Machine Intell. 23 (8), 800–810.

Dony, R.D., Wesolkowski, S.B. 1999. Edge detection on color images using RGB vector angles. In: Proc. 1999 IEEE Canadian Conf. Electrical and Computer Engineering. Edmonton, Alberta, Canada May 9–12, pp. 687–692.

Dupuis, A., Vasseur, P., 2006. Image segmentation by cue selection and integration. Image Vision Comput. 24 (10), 1053–1064, October.

Huang, K., Wang, Q., Wu, Z., 2006. Natural color image enhancement and evaluation algorithm based on human visual system. Computer Vision Image Understand. 103 (1), 52–63, July.

Kato, Z., Pong, T.C., 2006. A Markov random field image segmentation model for color textured images. Image Vision Comput. 24 (10), 1103–1114, October.

Martin, D., Fowlkes, C., Tal, D., Malik, J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proc. 8th Internat. Conf. on Computer Vision, vol. 2. pp. 416–423.

Megawave Homepage: <http://www.cmla.ens-cachan.fr/Cmla/Megawave/>, 2006 (accessed 14.9.06).

Mittelheusser, Gangolf, Kruggel, Frithjof, 1995. Fast Segmentation of Brain Magnetic Resonance Tomograms. In: CVRMed'95: First International Conference on Computer Vision, Virtual reality and Robotics in Medicine, INRIA, Nizza. Springer Verlag. April.

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

Priese, L., Rehrmann, V., 1993. A fast hybrid color segmentation method. DAGM-Symposium, 297–304.

Rehrmann, V., Priese, L., 1998. Fast and robust segmentation of natural color scenes. ACCV 1, 598–606.

Schneider, M.K., Fieguth, P.W., Karl, W.C., Willsky, A.S., 2000. Multiscale methods for the segmentation and reconstruction of signals and images. IEEE Trans. Image Process. 9 (3), 456–468, March.

Tilton, J.C., D-dimensional formulation and implementation of recursive hierarchical segmentation, Disclosure of Invention and New Technology: NASA Case No. GSC 15199-1, May 26, 2006.

Trémeau, A., Colantoni, P., 2000. Regions adjacency graph applied to color image segmentation. IEEE Trans. Image Process. 9 (4), 735–744.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Machine Intell. 13, 583–598.

Wu, Z., Leahy, R., 1993. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE Trans. Pattern Anal. Machine Intell. 15 (11), 1101–1113, Nov.