A FRAMEWORK FOR THE EFFICIENT SEGMENTATION OF LARGE-FORMAT COLOR IMAGES

Vasileios Mezaris, Ioannis Kompatsiaris, and Michael G. Strintzis

Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki 54006, Greece

Informatics and Telematics Institute
1st Km Thermi-Panorama Rd,
Thessaloniki 57001, Greece

ABSTRACT

In this paper, a novel approach to large-format image segmentation is presented, focused on usage in content-based multimedia applications. The proposed framework aims at facilitating the time-efficient segmentation of large-format images while maintaining the high perceptual quality of the segmentation result. For this to be achieved, the employed segmentation algorithm is applied to reduced versions of the large-format images, in order to speed-up its execution, resulting in a coarse-grained segmentation mask. The final fine-grained segmentation mask is produced by an enhancement stage that involves partial reclassification of the pixels of the original image using a Bayes classifier. As shown by experimental evaluation, this novel scheme provides fast segmentation with high perceptual segmentation quality.

1. INTRODUCTION

In recent years, the proliferation of digital media has established the need for the development of tools for the efficient representation, access and retrieval of visual information. While several approaches have been proposed to address these issues, most recent approaches rely on the analysis of the content of the medium in semantic objects. This is true both for still image manipulation (image indexing [1, 2], region-of-interest coding) and for video coding and indexing. The cornerstone of any such content-based application is the segmentation algorithm.

While several segmentation algorithms for use in multimedia applications have been presented, most of them handle mainly images of relatively small dimensions, e.g. 192×128 pixels. This is due to the fact that their computational complexity is at best proportional to the number of pixels of the image, making the segmentation of a large-format image particularly time-consuming. Since large-format images are becoming increasingly popular, time-efficient methods for their segmentation become essential. For this reason, a novel framework for the fast segmentation of large-format images (Fig. 1), that can be combined with various segmentation algorithms, is proposed in this paper.

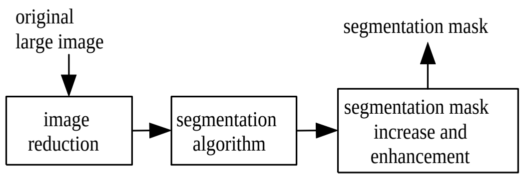

Fig. 1. Large-format image segmentation framework.

The paper is organized as follows: in section 2, the segmentation algorithm used in our experiments is presented. In section 3, the framework for the fast segmentation of large-format images is developed. Section 4 contains an experimental evaluation of the developed methods, and finally, conclusions are drawn in section 5.

2. COLOR IMAGE SEGMENTATION

The segmentation algorithm presented (Fig. 2) is based on a novel variant of the K-Means-with-connectivity-constraint algorithm (KMCC) [3, 4], a member of the popular K-Means family. The KMCC algorithm produces connected regions by taking into account the position of each pixel as well as its color and texture features.

The first step of the segmentation algorithm is the calculation of the color and texture features to be used for pixel classification. The color features used are the three intensity coordinates of the CIE L*a*b* color space; consequently the intensity feature vector of pixel $\mathbf{p}$, $I(\mathbf{p})$ is defined as

$$I(\mathbf{p}) = [I_L(\mathbf{p}), I_a(\mathbf{p}), I_b(\mathbf{p})]^T$$

In order to calculate a texture feature vector $T(\mathbf{p})$ for pixel $\mathbf{p}$, the 2-D fast iterative scheme for performing the Discrete

This work was supported by the European HISCORE IST project. The assistance of COST211 quat is also greatfully acknowledged.

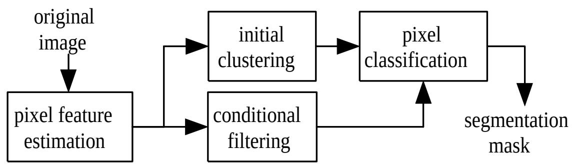

Fig. 2. Overview of the segmentation algorithm.

Wavelet Frames (DWF) decomposition, presented in [5], is used. The employed decomposition filter bank is based on the lowpass Haar filter, that was shown in [5] to be a suitable choice. For images of relatively small dimensions, e.g. 150×100 pixels, a 2-D DWF decomposition of two levels has been chosen ($L_d = 2$), whereas for the direct application of the algorithm to large-format images, four levels of decomposition were used instead of two, for effective texture characterization.

The pixel feature estimation is followed by an initial clustering procedure, in order to compute the initial values required by the KMCC algorithm. For that, the image is broken down to square non-overlapping blocks of dimension $f \times f$ and an intensity feature vector $I(b_l)$ and a texture feature vector $T(b_l)$ are assigned to each block. The number of regions of the image is initially estimated by applying a variant of the maximin algorithm to the blocks, followed by the application of a simple K-Means algorithm to them, K being equal to the estimated number of regions. This is followed by the application of a recursive four-connectivity component labelling algorithm [6], so that a total of K' connected regions s_k are identified. Their intensity, texture and spatial centers, $I(s_k)$, $T(s_k)$ and $S(s_k)$ respectively, are then calculated as the mean values of the features of the pixels belonging to the blocks assigned to each region.

In the conditional filtering step, a moving average filter alters the intensity features in those parts of the image where intensity fluctuations are particularly pronounced, since in these parts the KMCC algorithm generally does not perform well. The decision of whether the filter should be applied to a particular pixel $\mathbf{p}$ is made by evaluating the norm of its texture feature vector $T(\mathbf{p})$; the filter is not applied if that norm is below a threshold T_{th}. The output of the conditional filtering module can be expressed as:

$$J(\mathbf{p}) = \begin{cases} I(\mathbf{p}) & \text{if } \|T(\mathbf{p})\| < T_{th} \\ \frac{1}{f^2}\sum_{m=1}^{f^2} I(\mathbf{p}_m) & \text{if } \|T(\mathbf{p})\| \geq T_{th} \end{cases} \quad (1)$$

$$T_{th} = \max\{0.65 \cdot T_{max}, 14\} \quad (2)$$

where T_{max} is the maximum value of the norm $\|T(\mathbf{p})\|$ in the image.

Finally, pixels are classified into regions by a variant of the KMCC algorithm, using a generalized distance of a pixel $\mathbf{p}$ from a region s_k defined as:

$$D(\mathbf{p}, s_k) =$$
$$\|J(\mathbf{p}) - J(s_k)\| + \|T(\mathbf{p}) - T(s_k)\| + \lambda\frac{\bar{A}}{A_k}\|\mathbf{p} - S(s_k)\|$$

where $\|J(\mathbf{p}) - J(s_k)\|$, $\|T(\mathbf{p}) - T(s_k)\|$ and $\|\mathbf{p} - S(s_k)\|$ are the Euclidean distances of the intensity, texture and spatial feature vectors respectively, A_k is the area of region s_k, $\bar{A}$ is the average area of all regions and λ is a regularization parameter. The KMCC algorithm features splitting of non-connected regions and merging of neighboring regions with similar intensity or texture centers. The region centers are recalculated on every iteration as the mean values of the intensity, texture and spatial features of the pixels assigned to each region. Centers corresponding to regions that fall below a size threshold $th_{size} = 0.75\%$ of the image area, are omitted. Following the convergence of the KMCC algorithm, any regions smaller than th_{size} are eliminated by being forced to merge with one of their neighboring regions.

3. FAST LARGE-FORMAT IMAGE SEGMENTATION FRAMEWORK

The approach presented in the previous section is considerably fast when the algorithm is applied to images of relatively small dimensions, e.g. 150×100 pixels, but its time efficiency degrades quickly as the image size increases. In order to provide a more efficient scheme for the segmentation of large images, one could take advantage of a reasonable assumption already made in the previous section, namely that regions falling below a size threshold $th_{size} = 0.75\%$ of the total image area are insignificant. For relatively large images, this threshold corresponds to a large number of pixels. This reveals the potential of applying the segmentation algorithm to images reduced by a factor of $R > 1$.

A necessary condition for all significant objects to be detectible in the reduced image is that the size threshold for the reduced image, expressed as the minimum number of pixels, be much greater than one; thus,

$$\frac{th_{size} \cdot y_{max} \cdot x_{max}}{R^2} \gg 1,$$

where x_{max}, y_{max} are the original image dimensions.

Applying the employed segmentation algorithm to a reduced image improves the time efficiency of the segmentation process, but does so at the expense of the quality of the segmentation result, since edges between objects are crudely approximated by piecewise linear segments, lowering the perceptual quality of the result (fig. 4c). To alleviate this problem, the use of the Bayes classifier for the reclassification of pixels belonging to blocks on edges between

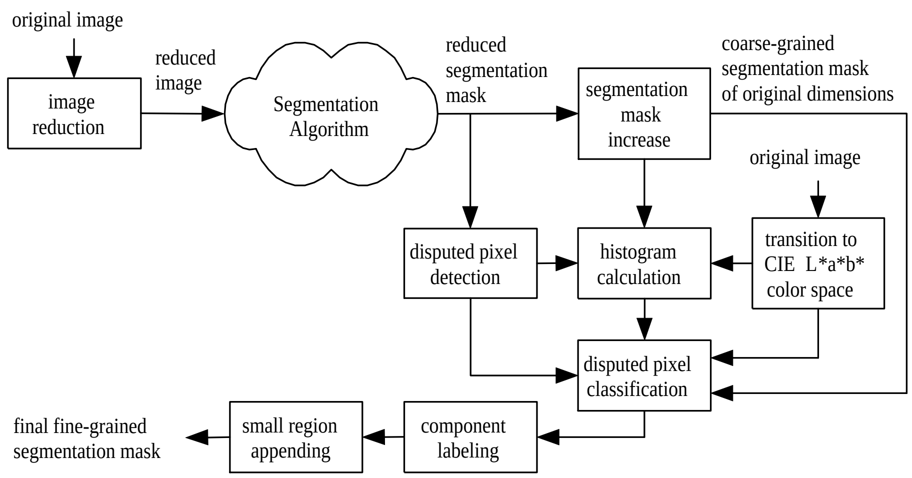

Fig. 3. Block diagram of the proposed framework for fast segmentation of large-format color images.

regions is proposed. If a block, assigned to one region, is neighboring to blocks of Γ other regions, $\Gamma \neq 0$, the assignments of all pixels of the original image represented by that block must be reevaluated, since each of them may belong to any one of the possible $\Gamma + 1$ regions. In this way, G sets $g_i^p, i = 1, \ldots, G$ of disputed pixels are formed, each associated with a different set $g_i^s, i = 1, \ldots, G$ of possible regions.

The reclassification of the disputed pixels is then performed using their intensity values in the CIE L*a*b color space, as follows: Let ω_k be the class of pixel intensities of region s_k, not including the pixel intensities of any disputed pixels :

$$\omega_k = \{I(\mathbf{p}), \mathbf{p} \in s_k, \mathbf{p} \notin g_i^p, i = 1, \ldots, G\}$$

According to the Bayes classifier for minimum classification error, a disputed pixel $\mathbf{p}$, $\mathbf{p} \in g_i^p$, is assigned to region s_k if

$$p(\omega_k | I(\mathbf{p})) > p(\omega_q | I(\mathbf{p})), \;\; \forall s_k, s_q \in g_i^s, k \neq q, \qquad (3)$$

Using the Bayes Theorem and assuming that among the pixels of group g_i^p the a priory probability of class ω_k, $p(\omega_k)$, $s_k \in g_i^s$, is equal for all regions $s_k \in g_i^s$, the classification criterion of equation (3) is simplified to: pixel $\mathbf{p}$, $\mathbf{p} \in g_i^p$, is assigned to region s_k if

$$p(I(\mathbf{p}) | \omega_k) > p(I(\mathbf{p}) | \omega_q), \;\; \forall s_k, s_q \in g_i^s, k \neq q$$

where

$$p(I(\mathbf{p}) | \omega_k) = \prod_{x \in \{L,a,b\}} hist_k^x(I_x(\mathbf{p})), \qquad (4)$$

and $hist_k^x$, $x \in \{L, a, b\}$, are the normalized histograms of region s_k, excluding any disputed pixels.

Since no connectivity constraint is enforced during the reclassification of the disputed pixels, the connectivity of the formed regions must be evaluated as soon as the reclassification is completed, using a four-connectivity component labelling algorithm [6]. This process is followed by the appending of any small regions: a small region s_{k_1} is appended to region $s_{k_2}, k_2 = 1, \ldots, K, k_2 \neq k_1$, for which the intensity distance

$$D_I(s_{k_1}, s_{k_2}) = \|I(s_{k_1}) - I(s_{k_2})\|$$

is minimum. The block diagram of the proposed framework is presented in figure 3.

4. EXPERIMENTAL RESULTS

The efficiency of the proposed fast large-format image segmentation framework of section 3 was evaluated by comparing its time efficiency and perceptual segmentation quality with two other segmentation schemes described in table 1. The time efficiency of the three aforementioned schemes was compared on an 800MHz Intel Pentium III PC, using a set of 100 images of 730×490 pixels from the Corel gallery [7]. The perceptual segmentation quality can be compared using the results of figure 4. These lead to the conclusion that the proposed framework is more efficient than the direct approach, since it delivers the same segmentation quality, with segmentation time significantly reduced. Applying the segmentation algorithm to reduced images without reclassification using the Bayes classifier is naturally even faster, but produces segmentation masks of lower perceptual quality. Another important observation is that the proposed framework makes no assumption about the employed segmentation algorithm, thus can be combined with a variety of segmentation algorithms.

Table 1. Average Segmentation Time (730×490 pixels)

Segmentation Scheme	Time (sec)
Direct application of the algorithm of section 2 to 730×490 pixel images	2494.28
Application of the segmentation algorithm of section 2 to reduced images (reduction factor $R = 8$)	18.92
Application of the fast large-format image segmentation framework of section 3 (reduction factor $R = 8$)	47.55

5. CONCLUSIONS

A methodology was presented for the time-efficient segmentation of large-format color images. The proposed framework is appropriate for use as part of a content-based multimedia application, such as image querying by example, or for defining regions of interest for content-based coding of still images, in the context of the JPEG2000 standard.

6. REFERENCES

[1] H. Greenspan S. Belongie, C. Carson and J. Malik, "Color- and texture-based image segmentation using EM and its application to content-based image retrieval," in *Proc. Sixth International Conference on Computer Vision*, January 1998.

[2] N. V. Boulgouris, I. Kompatsiaris, V. Mezaris, and M. G. Strintzis, "Content-based Watermarking for Indexing Using Robust Segmentation," in *Proc. Workshop on Image Analysis For Multimedia Interactive Services*, Tampere, Finland, May 2001.

[3] I. Kompatsiaris and M. G. Strintzis, "Content-Based Representation of Colour Image Sequences," in *IEEE International Conference on Acoustics, Speech, and Signal Processing 2001*, USA, May 7-11 2000.

[4] I. Kompatsiaris and M. G. Strintzis, "Spatiotemporal Segmentation and Tracking of Objects for Visualization of Videoconference Image Sequences," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 8, December 2000.

[5] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. on Image Processing*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.

[6] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, McGraw-Hill International Editions, 1995.

[7] *Corel stock photo library*, Corel Corp., Ontario, Canada.

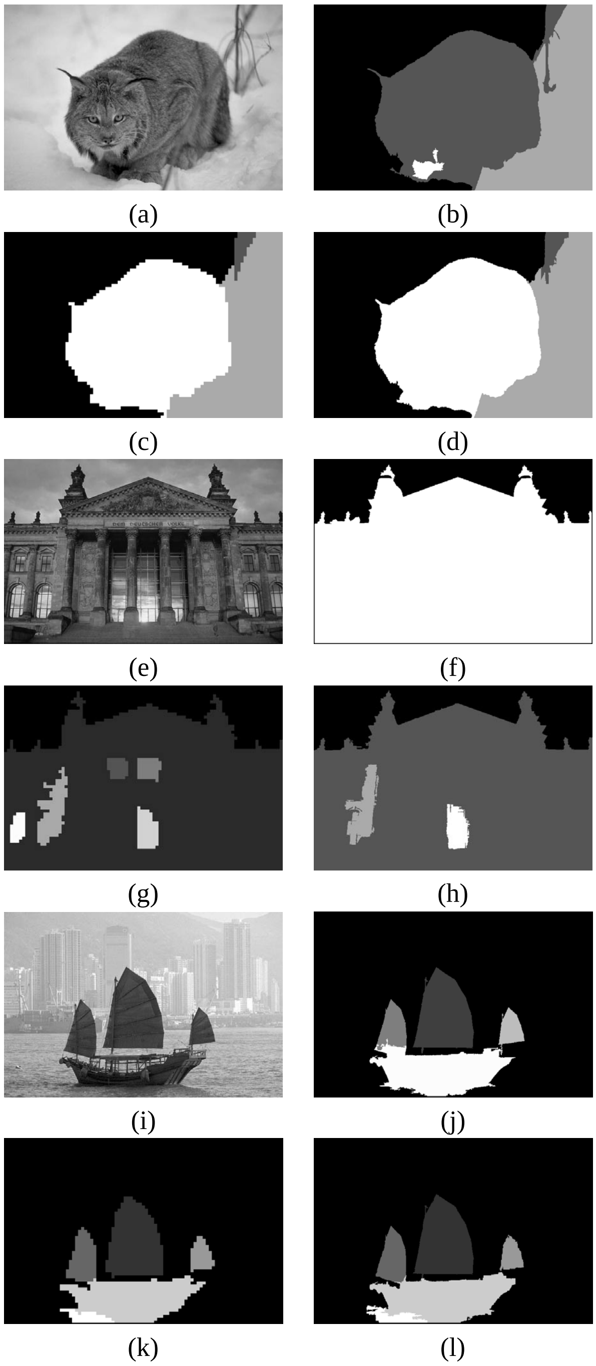

Fig. 4. (a),(e)&(i) Original images of 730×490 pixels. (b),(f)&(j) Results of the direct application of the segmentation algorithm of section 2. (c),(g)&(k) Results of the application of the same algorithm on reduced images (reduction factor $R = 8$). Edges between objects have been approximated by piecewise linear segments. (d),(h)&(l) Results of the large-format image segmentation framework.