STATISTICAL MULTISCALE IMAGE SEGMENTATION VIA ALPHA-STABLE MODELING

Tao Wan   Nishan Canagarajah   Alin Achim

Department of Electrical and Electronic Engineering
University of Bristol, Bristol, BS8 1UB, UK
{Tao.Wan, Nishan.Canagarajah, Alin.Achim}@bristol.ac.uk

ABSTRACT

This paper presents a new statistical image segmentation algorithm, in which the texture features are modeled by Symmetric Alpha-Stable (SαS) distributions. These features are efficiently combined with the dominant color feature to perform automatic segmentation. First, the image is roughly segmented into textured and nontextured regions using the Dual-Tree Complex Wavelet Transform (DT-CWT) with the subband coefficients modeled as SαS random variables. A multiscale segmentation is then applied to the resulting regions, according to the local texture characteristics. Finally, a novel statistical region merging algorithm is introduced by measuring the Kullback-Leibler distance (KLD) between estimated SαS models for the neighboring segments. Experiments show that our algorithm achieves superior segmentation results in comparison with existing state-of-the-art image segmentation algorithms.

Index Terms— multiscale image segmentation, statistical modeling, wavelet transform, SαS distributions, KLD

1. INTRODUCTION

Image segmentation is a fundamental problem in image processing and analysis. It provides a partitioning of the image in regions that should represent meaningful objects. In recent years, many authors have applied statistical techniques combined with wavelet transform for image segmentation [1][2]. These approaches have improved the segmentation results of different image modalities. In previous work [3], we have proposed a multiscale image segmentation algorithm based on dominant color and homogeneous texture features that are efficiently combined to perform the automatic segmentation. This paper presents alternative approaches to the texture feature extraction and the region merging components of the algorithm in [3]. The texture feature extraction is improved through modeling of the marginal distribution of wavelet coefficients via symmetric alpha-stable distributions, while the region merging is based on the Kullback-Leibler distance as similarity metric between two neighboring segments.

Recently, Achim *et al.*[4] have shown that successful statistical image processing algorithms can be developed if they take into consideration the actual heavy-tailed behavior of most real life signals. Specifically, they have shown that wavelet decomposition coefficients of images are best modeled by symmetric alpha-stable distributions, a family of heavy-tailed densities. In our work, the texture segmentation process integrates the Dual-Tree Complex Wavelet Transform [5] and alpha-stable statistical modeling [4] to characterize wavelet coefficients of natural images. The former provides near shift invariance and good directional selectivity compared to the standard wavelet transform, while the later has been shown to be a good approximation for the marginal density of the coefficients at a particular subband. In the region merging stage, the segment similarity is measured by the Kullback-Leibler distance between two SαS models corresponding to the adjacent segments. The key to the success of the proposed algorithm is that the statistical modeling techniques and low-level features of color and texture are integrated into a single image segmentation framework to achieve precise and robust segmentation.

The structure of the paper is as follows: the texture segmentation using DT-CWT and SαS is described in Section 2. The multiscale image segmentation algorithm is discussed in Section 3. Our proposed statistical region merging approach and experimental results are presented in Section 4 and Section 5 respectively. Finally, conclusions are summarized in Section 6.

2. TEXTURE SEGMENTATION USING DT-CWT AND SαS

Statistical modeling is much easier to perform in a suitable transform space where simple models with a small number of parameters can describe the data, rather than on the original image pixel values. Wavelets have emerged as an effective

Tao Wan was supported by an Overseas Research Students Award (ORSAS).

tool to analyze texture information as they provide a natural partition of the image spectrum into multiscale and oriented subbands. In this work,we use a three-scale DT-CWT [5] with six orientations, which is able to provide approximate shift invariance and directionally selective filters while preserving the usual properties of perfect reconstruction and computational efficiency. The marginal distribution of the subband coefficients is well represented by adaptively varying two parameters of SαS distributions. SαS distributions are best defined by their characteristic functions due to lack of a compact analytical expression for their probability density function:

$$\varphi(\omega) = \exp(j\delta\omega - \gamma|\omega|^{\alpha}) \tag{1}$$

where $\alpha(0 < \alpha \leq 2)$ is the characteristic exponent, $\delta(-\infty < \delta < \infty)$ is the location parameter, and $\gamma(\gamma > 0)$ is the dispersion of the distribution, similar to the variance of the Gaussian distribution. Since our further developments are in the framework of wavelet analysis, in the following we will assume that $\delta = 0$. Parameters α and γ can be estimated through the approach proposed in [6]. In our work, the above estimation is implemented in a square-shaped neighborhood of size 7×7 around each reference coefficient. Therefore, the texture feature value $T(x, y)$ at the pixel location (x, y) is defined as:

$$T(x, y) = \{\alpha_i(x, y), \gamma_i(x, y)\} \quad i = 1, 2, 3, ..., 18 \tag{2}$$

where i is the index of i^{th} subband.

In order to obtain a uniform characterization of texture, median filtering is employed on $T(x, y)$ within each subband to filter out the texture associated with transition between regions. Finally, a two-cluster K-means algorithm is used to classify the pixels to textured and non-textured regions with the high values of $T(x, y)$ assigned to textured regions and low values assigned to non-textured regions. A pixel is then classified as textured if the proportion of the number of the subbands belonging to the textured region is above a threshold P. In our experiments, a suggested value for the threshold is $P = 0.5$ for the color images. Compared with [3], the threshold can be adjusted to the type of the images. This property is useful for the proposed algorithm which can handle not only natural images, but also multimodal images. In Fig.1, we compare the proposed approach with simple methods that only use DT-CWT or Gabor coefficients as in [3]. From the figure it can be seen that the new approach provides more accurate texture segmentation with smoother contours.

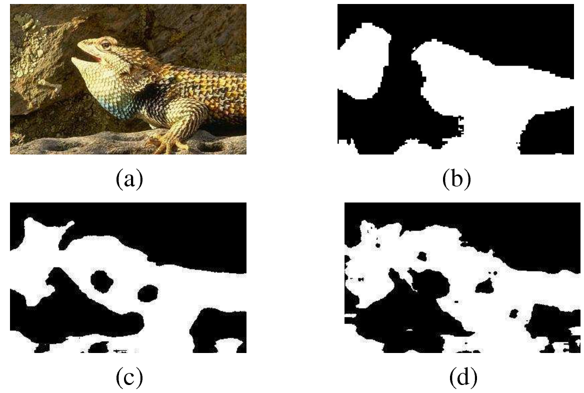


Fig. 1. Texture segmentation (a) Original image. (b) Texture map using SαS. (c) Texture map using DT-CWT. (d) Texture map using Gabor decomposition. White regions correspond textured regions, and black regions correspond nontextured regions.

3. MULTISCALE IMAGE SEGMENTATION

The textured and nontextured regions are further segmented into relatively small and homogeneous regions while retaining the boundaries between the two regions. The dominant colors are first extracted based on Peer Group Filtering (PGF) [7] and the Generalized Lloyd Algorithm [8]. Then, the JSEG algorithm proposed by Deng *et al.* in [7] is used to minimize the cost associated with partitioning an image at different scales. A bigger window size is used for high scales, which are useful for detecting texture boundaries, while lower scales are employed in order to localize the intensity of color edges. It is reasonable to apply the lower scales to the nontextured region, which has a more or less homogeneous texture, while higher scales are adopted for the textured region to find the texture boundaries. In contrast with JSEG, which doesn't take into account the local texture difference between the image regions, the strength of this approach is that we are able to apply the multiscale segmentation simultaneously to the same image according to the local texture characteristics.

However, the current boundary locations between textured and nontextured regions are not the actual boundaries due to the fact that K-means clustering can only segment the image into rough regions. Moveover, multiscale segmentation provides accurate results only within the textured and nontextured regions. Consequently, a boundary refinement step is employed to adjust the boundaries between the two regions. A pixel is assigned to the neighbor class that has the minimum D value using the following function:

$$D = Dist(C^0, C^i) + a(S_4^i - D_4^i) + b(S_8^i - D_8^i) \tag{3}$$

Where $Dist$ refers to the Euclidean distance measure, C^0 and C^i are the dominant color vectors of the current pixel and its i^{th} neighbor segment, S_4^i and S_8^i are the numbers of 4- and 8-neighbor pixels belonging to the ith segment, while D_4^i and D_8^i are the numbers of 4- and 8-neighbor pixels belonging to the different classes of the i^{th} segment. a and b represent the strength of the spatial constraint. Specifically, as a and b increase, a pixel is more likely to belong to the class to which many of its neighbors belong. Thus region boundary smoothness is achieved. The influence of a and b on the boundary

(a) (b) (c)

Fig. 2. Boundary refinement results (a) $a = 0.8, b = 0.0$. (b) $a = 0.0, b = 0.8$. (c) $a = 1.0, b = 0.8$.

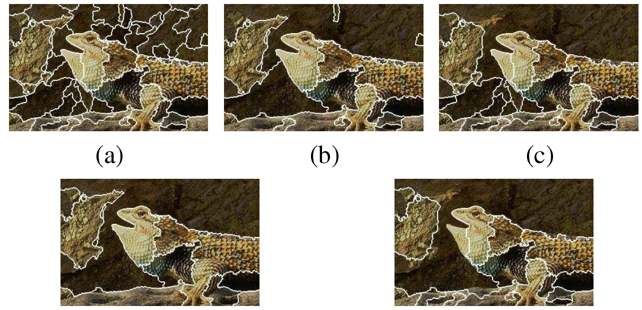
refinement procedure is shown in Fig.2. The result in Fig.2(c) that uses higher values of a and b has smoother boundaries compared with Fig.2(a) and 2(b).

4. STATISTICAL REGION MERGING

In general, the result of applying the algorithm described in the previous sections leads to over-segmentation. A statistical region merging method is implemented by using SαS distributions to appropriately model wavelet coefficients within the segmented regions. In this work, the regions are classified into two categories. The segments with more than 80% of their pixels belonging to the nontextured areas are categorized as nontextured segments, and the remaining segments are classified as textured segments. Therefore, segmented regions are considered individually rather than globally.

A corresponding merging criterion is provided for each category. The main difference lies in the way of feature extraction in the regions. Nontextured segments are merged based on their dominant color similarity. To achieve this, the Euclidean distance of the color histograms extracted from the neighboring nontextured segments is calculated. For textured segments, region similarity is measured using statistical model parameters followed by computing the Kullback-Leibler distance (KLD).

In [9], the authors introduced a statistical framework for texture image retrieval where the marginal density of coefficients is approximated by symmetric α-stable distributions, and texture similarity is measured by means of the Kullback-Leibler distance between model parameters. Inspired by their work, we model subband complex wavelet coefficients in the textured regions independently using SαS densities, where model parameters can be obtained through the method addressed in [6]. Therefore, the characteristics of the region can be completely defined via two parameters α and γ. There is no closed-form expression for the KLD between two general SαS distributions, which are not Cauchy or Gaussian. By applying the KLD on the normalized version of the SαS characteristic function, we expect to obtain good similarity measurement. The KLD between two adjacent textured segments



(a) (b) (c)

(d) (e)

Fig. 3. Region merging results (a) Result of the multiscale segmentation step. (b) Region merging using KLD. (c) Region merging using Euclidean distance. (d) Final merging result using KLD. (e) Final merging result using Euclidean distance.

is given by:

$$KLD(s_1, s_2) = \frac{1}{18} \sum_{j=1}^{18} \left(\ln\left(\frac{c_1^j}{c_2^j}\right) - \frac{1}{\alpha_1^j} + \frac{2\gamma_2^j \Gamma\left(\frac{\alpha_2^j+1}{\alpha_1^j}\right)}{c_1^j \alpha_1^j \gamma_1^{\frac{\alpha_2^j+1}{\alpha_1^j}}} \right) \tag{4}$$

$$c_i = \frac{2\Gamma\left(\frac{1}{\alpha_i}\right)}{\alpha_i \gamma_i^{1/\alpha_i}} \quad i = 1, 2 \tag{5}$$

where $\Gamma(\cdot)$ is the Gamma function, s_1 and s_2 are the adjoining textured segments, and j denotes the index of the wavelet subband. The pair of regions with the minimum distance is merged until a maximum threshold of the distance is reached. Compared to the previous work [3] in which the segments are classed into three categories, our two-category method offers comparable results with reduction in computational complexity. Fig.3(b) and 3(c) show the merging results after few iterations of the region merging algorithm using both KLD and Euclidean distance. Both images contain the same number of segmented regions. In addition, Fig.3(d) and 3(e) show the final segmentation at the end of the region merging process obtained using KLD and Euclidean distance respectively. Clearly, the KLD provides better results than Euclidean distance due to the fact that the statistical measures of similarity are more accurate than typically norm-based distances in terms of human visual perception.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

The segmentation algorithm has been evaluated on a variety of natural images. Fig.4 shows the segmentation results obtained using four different methods, including the multiscale color-texture segmentation [3], JSEG [7], the watershed algorithm [10] and our proposed algorithm. In addition, the figure also includes the ground truth images from Berkeley dataset

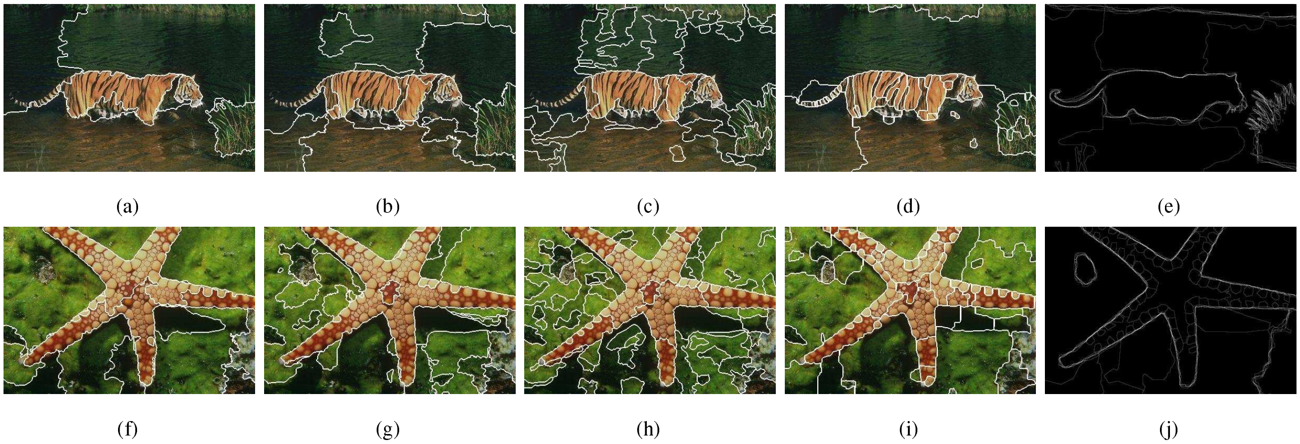


Fig. 4. Segmentation results using "tiger" and "starfish" images. From left to right: statistical multiscale segmentation, multiscale color-texture segmentation [3], JSEG [7], watershed algorithm [10], and hand-labeled ground truth image segmentation form Berkeley dataset [11].

[11]. On inspecting our results (Fig. 4 (a) and (f)), it is clear that the tiger and the starfish, which are salient objects in the two images, are better segmented than by JSEG. The watershed is also able to accurately detect contours but still exhibits an over-segmentation problem within salient objects. Experiments also indicate that our approach provides superior segmentation results with more accurate and smoother boundaries in comparing with our previous method [3] shown in Fig.4(b) and 4(g). The improvements lie in the better texture feature extraction, as well as the similarity measurement via the Kullback-Leibler distance. The former helps in obtaining more accurate initial texture segmentation results, while the later enhances the statistical region merging procedure.

6. CONCLUSIONS AND FUTURE WORK

This paper has demonstrated that robust and meaningful image segmentation can be achieved by integrating SαS modeling with color-texture features that are widely used in image processing. The local statistical characteristics have been taken into account in the texture segmentation process with the adaptive threshold. Consequently, this approach can be interesting for the segmentation of multimodal images. The main contribution of this work is that it provides an accurate automatic image segmentation through a single framework which efficiently combines the statistical techniques, multiscale analysis and low level features of color and texture. Future work will concentrate on combining the segmentation results obtained to image fusion applications, and extending the method to content-based video analysis.

7. REFERENCES

[1] D. Boukerroui, O. Basset, and A. Baskurt, "Multiresolution adaptive image segmentation based on global and local statistics," in *Proc. IEEE Conf. Image Process.*, 1999, vol. 1, pp. 358–361.

[2] C. R. Jung, "Combining wavelets and watersheds for robust multiscale image segmentation," *Image and Vision Computing*, vol. 25, no. 1, pp. 24–33, 2007.

[3] T. Wan, N. Canagarajah, and A. Achim, "Multiscale color-texture image segmentation with adaptive region merging," in *Proc. IEEE Conf. Acoustics, Speech, and Signal Process.*, Hawaii, Apr. 2007.

[4] A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Tran. Geosci. and Remote Sensing*, vol. 41, pp. 1773–1784, Aug. 2003.

[5] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering," *Appl. Compt. Harmon. Anal.*, vol. 10, pp. 234–253, May 2001.

[6] X. Ma and C. L. Nikias, "Parameter estimation and blind channel identification in impulsive signal environment," *IEEE Tran. Sign. Proc.*, vol. 43, pp. 2884–2897, Dec. 1995.

[7] Y. Deng, B. S. Manjunath, and H. Shin, "Color image segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 1999, vol. 2, pp. 446–451.

[8] A. Gersho and R. M. Gray, *Vector quantization and signal compression*, Kluwer Academic Publisher, Norwell, Mss., 1993.

[9] G. Tzagkarakis and P. Tsakalides, "A statistical approach to texture image retrieval via alpha-stable modeling of wavelet decompositions," in *Proc. 5th Int. Workshop on Image Analysis for Multimedia Interactive Services*, Apr. 2004.

[10] R. J. O'Callaghan and D. R. Bull, "Combined morphological-spectral unsupervised image segmentation," *IEEE Tran. Image process.*, vol. 14, no. 1, pp. 49–62, Jan. 2005.

[11] *The Berkeley Segmentation Dataset Benchmark*, University of California, Berkeley, CA, http://www.eecs.berkeley.edu/research/projects/cs/vision/grouping/segbench/ edition, Jul. 2003.