Natural Scene Segmentation Based on Information Fusion and Homogeneity Property

Heng-Da Cheng[1], Manasi Datar[2], Wen Ju[3]

[1]hengda.cheng@usu.edu
[2]manasi.data@gmail.com [3]wenju@cc.usu.edu
Computer Science Department, Utah State University, Logan, UT 84322-4205

Abstract

This paper presents a novel approach to natural scene segmentation. It uses both color and texture features in cooperation to provide comprehensive knowledge about every pixel in the image. A novel scheme for the collection of training samples, based on homogeneity, is proposed. Natural scene segmentation is carried out using a two-stage hierarchical self-organizing map (HSOM). The proposed method confirms that the sample selection based on homogeneity and the self-learning ability and adaptability of the HSOM, coupled with the information fusion mechanism, can lead to good segmentation result, which is validated by experiments on a variety of natural scene images.

Keywords: image segmentation, homogeneity, color feature, texture feature, information fusion, HSOM.

1. Introduction

Image segmentation is a process of dividing an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous. Due to the abundant information contained in color images, color image segmentation attracts more and more attention [1].

Recently, artificial neural network approach is being extensively studied. Self-organizing map (SOM) network, based on the idea of preserving the topology of the original input data set, was first proposed by Kohonen [2]. Unlike simple competitive learning methods in which only the winning neuron is allowed to learn, the neurons in the neighborhood of the winning neuron also participate in the learning process, leading to an ordered feature-mapping that can be exploited in many applications. The limitation of this method is that the final number of colors has to be specified a priori. Lampinen and Oja [3] have proposed a multi-layer SOM, HSOM. It forms arbitrarily complex clusters and produces clusters that better match the desired classes than the direct SOM algorithms.

Texture refers to a pattern of closely placed elements in such a manner that the pattern somehow repeats itself. It contains useful information and has long been an active area of research in the image processing community. Laws [4] developed a coherent set of "texture energy" masks which can be used to extract feature images for segmentation and classification.

In this paper, we propose a method for the segmentation of natural scenes by a two-stage HSOM. The input of this two-stage HSOM contains both color and texture features. The texture features are obtained by applying Laws' texture energy measures to the raw image. A novel approach for selecting training samples for the HSOM, based on homogeneity, is proposed. It ensures that the training dataset contains more pixels representing the diverse regions in the image, rather than those representing the homogeneous regions. Since the output of the HSOM is often an over-segmented image, the region-merging phase is implemented to achieve the final segmented image.

The rest of the paper is organized as follows. Section 2 describes the proposed segmentation method. Section 3 illustrates the experimental results. Section 4 draws conclusions.

2. The Proposed Method

The flowchart of the method proposed is shown in Fig. 1. The following subsections will explain each component in detail.

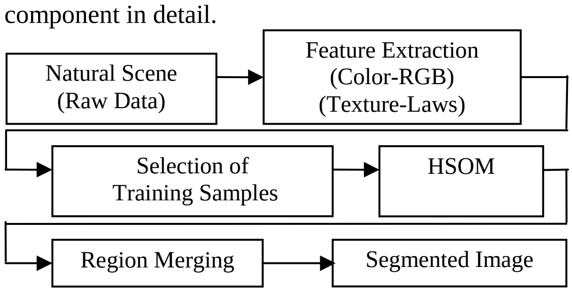


Fig. 1: The flowchart of the algorithm proposed

2.1. Feature Extraction

The objective of this stage is to represent each pixel as a vector of features suitable for inputting into the SOM for training and subsequent classification. For this research, each pixel is represented as a seven-dimensional vector:
$\{r, g, b, e5l5, e5s5, r5r5, l5s5\}$. The first three components of the feature vector are R, G, B values for each pixel in the original image. The next four components are obtained by applying the Laws' texture energy measures [4]. Fig. 2 depicts the four Laws' masks used in this paper.

-1	-2	0	2	1
-4	-8	0	8	4
-6	-12	0	12	6
-4	-8	0	8	4
-1	-2	0	2	1

L5E5

-1	0	2	0	-1
-2	0	4	0	-2
0	0	0	0	0
2	0	-4	0	2
1	0	-2	0	1

E5S5

1	-4	6	-4	1
-4	16	-24	16	-4
6	-24	36	-24	6
-4	16	-24	16	-4
1	-4	6	-4	1

R5R5

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

L5S5

Fig. 2: Four of Law's most successful masks

At the end of this stage, the image is represented by a matrix in which each row is a feature and each column is a pixel location. Assuming that the image has M rows and N columns, the input data for the SOM is a matrix of feature values for all the pixels in the image, formulated as $7 * (M * N)$ matrix.

2.2. Selection of Training Samples

The literature that discusses the selection of the samples for training the SOM is scarce. Random selection is the most commonly used. While random selection ensures an unbiased collection of training samples, it does not always provide the optimal set of training samples. In the case of image segmentation, the pixels around the boundary of the perceptual segments provide more information and should be used for training. Therefore, a novel approach for selecting training samples is proposed in this paper. The selection criterion is based on a definition of homogeneity proposed in [5].

2.2.1 Homogeneity Measure

Homogeneity is mainly related to the local information of an image and reflects how uniform an image region is. The more uniform the local region surrounding a pixel is, the larger the homogeneity value for the pixel. In [5], the homogeneity is denoted by β_{ij} for pixel location (i, j). However that homogeneity measure holds only for grayscale images. In order to be used in color image, we extend the concept to the domain of RGB images. Suppose β_{Rij}, β_{Gij} and β_{Bij} are the homogeneity measures calculated in the R, G and B color planes respectively, the homogeneity measure for the pixel location (i, j) in the RGB domain can be defined as:

$$\beta_{RGB_{ij}} = 0.33 \times \beta_{R_{ij}} + 0.33 \times \beta_{G_{ij}} + 0.33 \times \beta_{B_{ij}} \quad (1)$$

The non-homogeneity measure in the RGB domain can be calculated as:

$$\varphi_{RGB_{ij}} = 1 - \beta_{RGB_{ij}} \quad (2)$$

2.2.2 Selection Algorithm

We will use the non-homogeneity measure defined above in the RGB domain to obtain training samples. The proposed selection algorithm consists of following steps:

1. We define a location set Φ to contain the pixel locations of all training samples and initialize it to empty.
2. The average non-homogeneity value is calculated for the entire image as
$$\mu_{\varphi_{image}} = \frac{1}{MN} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \varphi_{RGB_{pq}} \quad (3)$$
3. The image is divided into blocks of arbitrary size $d*d$ (in this paper, d=15), and the local average nonhomogeneity value for each block t is calculated as
$$\mu_{\varphi_{block}} = \frac{1}{d^2} \sum_{p=i-\frac{d-1}{2}}^{i+\frac{d-1}{2}} \sum_{q=j-\frac{d-1}{2}}^{j+\frac{d-1}{2}} \varphi_{RGB_{pq}} \quad (4)$$
4. For each $d*d$ block t of the image, the number of pixels to be chosen for training is decided by the threshold
$$n_{training} = \begin{cases} (\mu_{\varphi_{block}} - \mu_{\varphi_{image}}) * d^2 & if\,(\mu_{\varphi_{block}} - \mu_{\varphi_{image}} > 0) \\ 10 & if\,(\mu_{\varphi_{block}} - \mu_{\varphi_{image}} <= 0) \end{cases} \quad (5)$$
5. $n_{training}$ pixel locations are then randomly selected from that block t and are added to the location set Φ.
6. Repeat steps 2-4 for all the blocks in the image.
7. The vectors corresponding to the locations in set Φ are then extracted from the SOM input matrix to form the final training set.

Selection of pixels based on the nonhomogeneity criterion ensures that the training dataset contains more pixels representing the diverse regions in the image, rather than those representing the homogeneous regions. Therefore the training dataset achieved in this way will carry more information of the image than the training dataset obtained by random selection and it will lead to better results of segmentation.

2.3. HSOM Training and Testing

In a conventional self-organizing map (SOM) network, the number of regions in the final segmented image depends on the number of neural units in the Kohonen layer, but it is very improbable that the number of regions in an image is known a priori. This significant shortcoming is overcome by implementing a hierarchical two-stage self-organizing network (HSOM) as a pattern classifier to group the output neurons into subsets, each of which corresponds to a discrete region [8]. In this paper, we use the same structure of HSOM as in [8] except for different number of neurons in both stages. There are 100 output neurons arranged in a 10 x 10 grid in the first stage and 20 neurons in the second stage. Especially, we use different approach to select training samples as discussed in section 2.2.

After undergoing the training process using the training dataset obtained in section 2.2, we can obtain the segmented image from the output of the HSOM.

2.4. Region Merging

The output of the HSOM is often an over-segmented image. Hence, the region-merging is implemented to combine regions that are similar to each other. The RGB color model is not suitable for measuring color difference. Hence, the image is converted into the CIE (*l*a*b**) color space, and the region merging process proposed in [6] is carried out. This phase generates the final segmentation result. The segmented image thus generated is more homogeneous within the regions and more disparity between them.

3. Experiment Results

The proposed approach has been tested on a variety of color images of natural scenes. Segmentation results, confirm that this approach is suitable for a wide range of images.

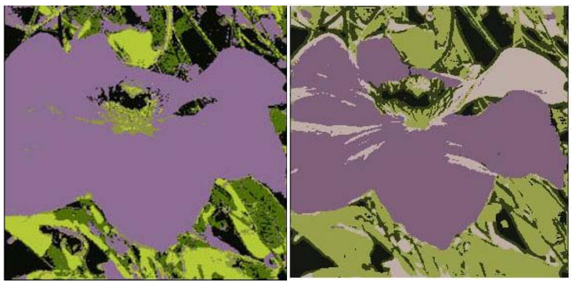
In order to demonstrate the performance of the proposed approach, we compare the results with other approaches.

[7] proposes a fuzzy homogeneity and scale space approach for color image segmentation. Fig. 3 shows the original image of “flower”. Fig.4 (a) is the segmentation result of [7] with 5 colors. Fig.4 (b) is the segmentation result of proposed method with 5 colors. In Fig. 4 (a) there are many pixels at the top and bottom of the picture that are misclassified as purple color. However, most of them are correctly classified in Fig. 4 (b). Also Fig. 4 (b) has more detail than Fig. 4 (a) while they have the same number of final colors.

The original “rose” image in Fig. 5 has vivid colors and is difficult to segment. [7] approximates it to six colors (Figure. 6 (a)), but [7] is still not able to provide the homogeneity evident in the segmentation resulting from the proposed approach (Figure. 6 (b)).The result of the proposed approach will be more valuable for the purpose of classification.

We also compare our results with the approach in [9], which uses a variational Bayesian framework for color image segmentation. Fig. 7 shows the original image of “sunset”. Fig.8 (a) is the segmentation result of [9] with 7 colors. Fig.8 (b) is the segmentation result of proposed method with 5 colors. In Fig. 8 (a) many pixels at the top of the picture are misclassified as leaves of the coconut tree while they are correctly classified in Fig. 8 (b). Also in Fig. 8 (b), the segmentation result is more homogeneous than in Fig. 8 (a).

Fig. 3: Original “flower” image.



(a) (b)

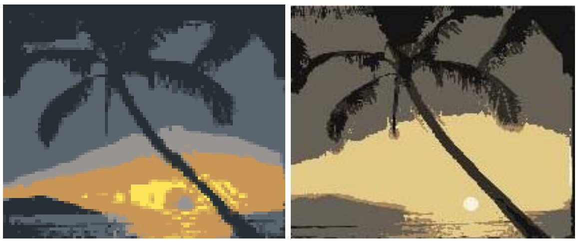
Fig. 4: (a) Segmentation result of [7] (5 colors). (b) The Segmentation using the proposed method (5 colors).

Fig. 5: Original "rose" image.

(a) (b)

Fig. 6: (a) Segmentation result of [7] (6 colors). (b) The Segmentation using the proposed method (6 colors).

Fig. 7: Original "sunset" image.



(a) (b)

Fig. 8: (a) Segmentation result of [9] (7 colors). (b) The Segmentation using the proposed method (5 colors).

The above comparisons demonstrate that the proposed approach performs the task of segmentation more effectively than those to be compared with.

4. Conclusion

The proposed approach based on the HSOM, uses both color and texture features in cooperation to provide comprehensive knowledge. Selection of training pixels based on the nonhomogeneity criterion ensures that the training dataset contains more pixels representing the diverse regions in the image, rather than those representing the homogeneous regions. The network detects the dominant color and texture features in a given color image. These are subsequently used to segment the image by pixel classification. Extensive experiments have shown that the proposed method is highly robust, flexible and effective. This approach will provide meaningful information by simplifying the image and identifying its constituent regions, thus helping interpretation of natural scene images.

5. References

[1] H. D. Cheng, X. H. Jiang, Y. Sun, and J. L. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, Vol. 34, pp. 2259-2281, 2001.

[2] T. Kohonen, "The self-organizing map," in *Proceedings of the IEEE*, ser. 9, Vol. 78, pp. 1464-1480, September 1990.

[3] J. Lampinen and E. Oja, "Clustering properties of hierarchical self-organizing maps,"*Journal of Mathematical Imaging and Vision*, Vol. 2, no. 2-3, pp. 261-272, November 1992.

[4] K. I. Laws, "Texture energy measures," in *Proccedings of Image Understanding Workshop*, pp. 47-51, November 1979.

[5] H. D. Cheng, M. Xue, and X. Shi, "Contrast enhancement based on a novel homogeneity measurement," *Pattern Recognition*, Vol. 36, no. 11, pp. 2687-2697, 2003.

[6] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Transactions on Image Processing*, Vol. 9, no. 12, pp. 2071-2082, December 2000.

[7] H. D. Cheng and J. Li, "Fuzzy homogeneity and scale-space approach to color image segmentation," *Pattern Recognition*, Vol. 36, no. 7, pp. 1545-1562, 2003.

[8] S. H. Ong, N. C. Yeo, K. H. Lee, Y. V. Venkatesh and D. M. Cao, "Segmentation of color images using a two-stage self-organizing network," *Image and Vision Computing*, Vol. 20, no. 4, pp. 279-289, 2002.

[9] N. Nasios and A. G. Bors,"A variational approach for color image segmentation," *Proceedings of the 17th International Conference on Pattern Recognition*, Vol. 1, pp. 680-683, 2004.