DOI: 10.3745/KIPSTB.2008.15-B.5.407

자기 조직화 기법을 활용한 컬러 영상 배경 영역 추출

신 현 경†

요 약

잡음이 심한 배경을 가진 영상 내부의 영역 분할 처리 과정은 해결하기 매우 어려운 문제로 인식되어 왔다. 그에 따라 이 문제를 해결하기 위한 기초적 방법론에 관한 연구 및 주어진 문제에 따라 실제적 적용을 위한 다양한 노력이 있어왔다. 본 논문에서는 영상 분할을 위한 새로운 접근법을 제시하는 것을 목적으로 하였다. 새로운 방법론으로서 기존의 관심 객체 분할의 반대인 배경 영역 분할이라는 새로운 관점을 연구의 중심으로 하였다. 기반 이론으로는 승자 독식 원리의 자기 학습 이론 알고리즘에서 특징 선택을 위한 자기 조직화를 분석하고 이를 문제 해결에 적용하였다. 실제적 영상 데이터를 통한 실험을 통해 배경 영역 분할을 적용한 영상 분할은 효과적으로 수행될 수 있음을 실험 결과로 제시해 보였다.

키워드 : 컬러 영상 분할, 배경 영상 분할, 특징 선택, 자기 조직화 사상

Background Segmentation in Color Image Using Self-Organizing Feature Selection

Shin Hyun Kyung†

ABSTRACT

Color segmentation is one of the most challenging problems in image processing especially in case of handling the images with cluttered background. Great amount of color segmentation methods have been developed and applied to real problems. In this paper, we suggest a new methodology. Our approach is focused on background extraction, as a complimentary operation to standard foreground object segmentation, using self-organizing feature selective property of unsupervised self-learning paradigm based on the competitive algorithm. The results of our studies show that background segmentation can be achievable in efficient manner.

Keywords : Color segmentation, Background segmentation, Feature selection and self organizing map

1. Introduction

Scanner manufacturing industry has need of a technology that enables for scanners to detect color regions in a document. Having color regions been detected, scanner can save input image with smaller file size without loss of perceptual information. One of the ideas is in using multi-layered imaging model. Outline of the process is the following: an input image will be divided by two layers: bi-tonal and color layer. In bi-tonal layer, black-and-white scanned image will be stored, and in color layer, the color scanned regions detected will be recorded. This is also an essential concept of design of MRC encoding which is under development in AdobeTM, XeroxTM and etc [1]. However, problems occur when a color document image contains non-homogeneous color background. In some cases a whole image area is considered as color region, which results in no benefit of reduction in file size, in other cases incorrectly localized color regions generate poor perceptual quality.

Due to ambiguity of definition and problem domain dependency, it is generally accepted that image segmentation is the most challenging task in image processing [2, 3]. For the last few decades, a great number of image segmentation methods have been emerged. In a survey paper [4], the authors discussed on recent updates of color segmentation which can be categorized by three methods: feature-based, image-domain based, and physics based. In survey papers [5, 6] the methods of performance evaluation on various image segmentation techniques are discussed.

† 종신회원 : 경원대학교 수학정보학과 조교수
논문접수 : 2008년 5월 26일
수 정 일 : 1차 2008년 7월 2일 2차, 2008년 7월 17일
심사완료 : 2008년 7월 18일

In this paper, in order for achieving color segmentation, we adopt an image-domain based method among the three basic categories aforementioned. The difference of our approach with other ones explained in section 2 is the following. Our study is focused on develop an algorithm that characterizes background regions in a document image. Consequently, by taking the information of regions of background detected as prior, results of foreground segmentation will be enhanced. Mathematically, image segmentation is not well-defined but over-determined problem as can be seen in many inverse problems. To make things worse, in practical applications, it is not unusual case when a given image contains severely cluttered background. By introducing a new perspective on definition of background, "background is widely connected exterior-like region in which pixel neighborhood has similarity in some sense of criteria", we can see a possibility of advanced image segmentation algorithm working on extensive range of images in the sense of complexity of background formation. In this paper, we address a question: "is there a good algorithm to identify the region of back-ground in color image?" Objective of the background seg-mentation is opposite to the typical color segmentation. Our assertion is that performance of color segmentation will be enhanced if it applies only to the complimentary region of the background identified.

Contents of this paper are organized as follows. In section 2, we investigated previous research works. In section 3, we describe the structural design of self-organization map applied on the framework of Kohonen network. In section 4, we discuss on the procedural point of self-organization learning algorithm. In section 5, experimental results for the study of this paper will be presented.

2. Previous Research Works

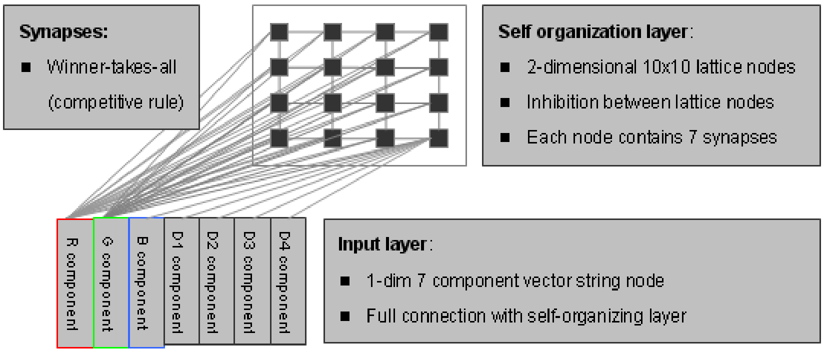
As a part of text information extraction project, document segmentation techniques have been addressed by several research groups. Previous research works are listed cate-gorically in aspect of their methods adopted. First of all, feature based method: Hase et al. [7] suggested a feature-based method using selective local color averaging and color space histogram analysis; Being specialized on low resolution image, Jia et al. [8] and Lim et al. [9] proposed 'tensor voting' method. Secondly, image domain based method: Ueda et al. [10] used a six-dimension vector consisted of RGB value and additional three color vector information. In the paper they tried to identify the regions of signature and seal imprints located in Japanese bank-check. Soria [11] adopted fuzzy integral operator on HIS color space in order to place the regions of official seals in color document image. Worring and Todoran [12] introduced a new clustering method exploiting spatial relation information. Thirdly, physics based method: Lucchese et al. [13] and Cheng et al. [14] investigated classical clustering techniques with K-means and C-means, respectively. Hammouche et al. [15] utilized statistical analysis on multi-dimensional local texture to introduce a co-occurrence matrix. Manjunath and Chellappa [16] improved Lakshamanan and Derin algorithm by applying unsupervised segmentation method based on random field model. As a new approach, Chakraborty and Duncan [17] adopted game theory in order to integrate classical region based and gradient based methods. Zhu et al. [18] used neural networks combined with wavelet methods. Bruce et al. [19] discussed about real-time mobile robot applications.

3. Design of Kohonen Networks

As an effort to resolve the questions addressed at the end of section 1, we adopt discriminative classifier property of self-learning model rather than generative one since we consider the segmentation problem as over-determined system. We applied the idea on the neural network frame-work among other statistical decision models. Between supervised and un-supervised learning paradigm, we selected the un-supervised since we realized that creation of training data for background region is not achievable. To be more specific, if our goal is to find the text regions in image, as an example of typical foreground segmentation, we can create training sample bitmap data containing various types of text region to teach neural network. However, construction of training sample for back-ground region is not simple. In order to incorporate with clustering technique for the image domain based segmentation, we adopted the competitive learning rule on the framework of two dimensional Kohonen self organization map [20, 21, 22] among various unsuper-vised learning rules.

In this section we describe implementation details on the Kohonen network. In structural point of view, the network consists of two neuron layers: input layer and two-dimen-sional self-organization layer. The input layer is in a form of one-dimensional string vector constituent of seven com-ponents. The self-organizing layer is constructed by ten-by-ten neurons. Refer to (Figure 1).

Like the standard Kohonen networks [23], each neuron in the lattice of self-organizing layer has the linking synapses which are fully connected to the input nodes and has inhibition relation with the neighborhood nodes. The concept of neighborhood defined in the space of



(Fig. 1) Structure of the Kohonen network adopted in this paper: self-organization layer is actually 10x10-lattice of nodes, and 7 components one dimensional string vector is used as input layer. For simplicity of drawing presentation, only 4x4-lattice of nodes is appeared.

lattice is topological. The learning algorithm adopted is the standard competitive learning rule.

For the clarity of notations used in the Kohonen map, our definitions of parameters are summarized in the <Table 1>:

〈Table 1〉 The parameters defined in Kohonen self-organization map

Name of parameter	Notation	Description
Input node	u_i	i = 0,⋯, 6.
Neuron in lattice	v_j (or v_{jk})	j = 0,⋯, 99 (or j,k = 0, ⋯, 9)
Weight on synapse	w_{jl}	j = 0,⋯, 99 and l = 0,⋯,6
Time constant 1	τ_1	Dynamic limiter to the Gaussian neighbor function
Time constant 2	τ_2	Dynamic limiter to the ordering phase
Learning rates	η	Limiter to synaptic weight update procedure

Internal mechanisms designed in the neural networks are as follow: initialization of synaptic weights, competition, cooperation, synaptic adaptation.

[Initialization] For initialization of synaptic weights w_{jk} the random values ranging 0 to 1 are assigned, where j is an index of lattice node, k is an index of input node.

[Competition] For competition stage, the winner node is chosen in the manner of competition process as shown in Eq.(1).

$$m = arg\ min(j)\ d(u,\ w_j), \tag{1}$$

where u is an input vector, and w a synaptic weight vector at j-th lattice node. The value m in Eq.(1) indicates that m-th node in the lattice is the winner node. The usual distance measure used is defined in Eq.(2) below.

$$d(u,\ w_j) = u^T \bullet w_j = \Sigma\ u_l\ w_{jl} \tag{2}$$

[Cooperation] For cooperation stage, inhibition process is performed with the non-stationary Gaussian neighborhood function (refer to Eq.(3)).

$$h_{j,i(x)}(n) = exp(-d^2_{i,j}/2\sigma^2(n)), \tag{3}$$

$$d^2_{i,j} = \|\ j - i\|^2, \tag{4}$$

$$\sigma\ (n) = \sigma_0\ exp(-n/\ \tau_1), \tag{5}$$

where i, j in Eq.(4) represent the indices of nodes in lattice. This implies that we used the topological ordering in the lattice with the Euclidean distance between nodes.

[Synaptic update] For synaptic update stage, for the sample vector x and the corresponding winner neuron index, i, the formula to update synaptic weight vector is in Eq.(6)-(7).

$$\triangle w_j(n) = \eta(n)\ \bullet\ h_{j,i(x)}(n)(x - w_j), \tag{6}$$

$$\eta(n) = \eta_0\ exp(-n/\ \tau_2), \tag{7}$$

where the number n indicates the iteration step for each node in lattice, i.e., j = 0, ⋯, 99.

As seen above, in order for the simulation there are several constant parameters to be set up. The actual values used in the simulation are summarized in <Table 2>.

〈Table 2〉 setting up the constant parameters

Parameter name	Notation	Value	Parameter name	Notation	Value
Time constant 1	τ_1	1000/log σ_0	Time constant 2	τ_2	1000
Learning rates	η	0.1	Neighboring const	σ_0	4

4. Discussion on Self-organization

In this section, acquisition of the sample data set for the self-organization adopted in this paper will be clarified. At each pixel of image, the RGB color channels and spatial connectivity information are selected to create a seven-dimensional vector. Refer to (Figure 1) for the complete list of each component defining the seven-dimensional vector. To be more specific, consider a color image, each pixel in the input image contains (R, G, B) channel values. If we limit ourselves to take RGB vector only as input feature selection, then it is a typical feature-based segmentation. For the purpose of performance enhancement, we included additional information for input feature selection. The additional information is the gradient of RGB values. Accordingly, the input vector for neural network is formed as [R, G, B, D1, D2, D3, D4]^T, where D indicates difference of RGB vectors with neighboring pixels. For convenience of presentation, we only show the seven-dimensional input vector. However, we also used higher dimensional input vector. For example, we used other statistical quantities such as local edge-map information - density of edges, histogram of edge density, and symmetry of edge directions.

Up to the current knowledge of the authors, background segmentation using neural networks is the first attempt. In the course of attempting the classical color segmentation methods, we observed that, especially feature-based approaches, the methods did not work well when the color image has non-uniform background. The idea of background segmentation came up when we tried to extract the color regions from the color document images.

Self-organization is the key part of the simulation. To show our Kohonen network works correctly, two-dimensional projection view of self-organizing feature map is presented. Refer to (Figure 2).

5. Experimental Results

[**Acquisition of data**] The data for this experiment is 23 test images provided by KodakTM for color document image processing. Training the Kohonen networks with 100x100 competition layer is computationally very heavy and takes about 45 seconds for the images with size of 2000 × 3000 in pixel. The numbers of iterations for the ordering and the convergence phase of training Kohonen are about 1000 and 10000 in average, respectively. Once the network trained, the prediction stage is straightforward and takes less than 100 milliseconds. There is a simple way to avoid time delay in learning, the off-line learning, which is out of the scope of this paper.

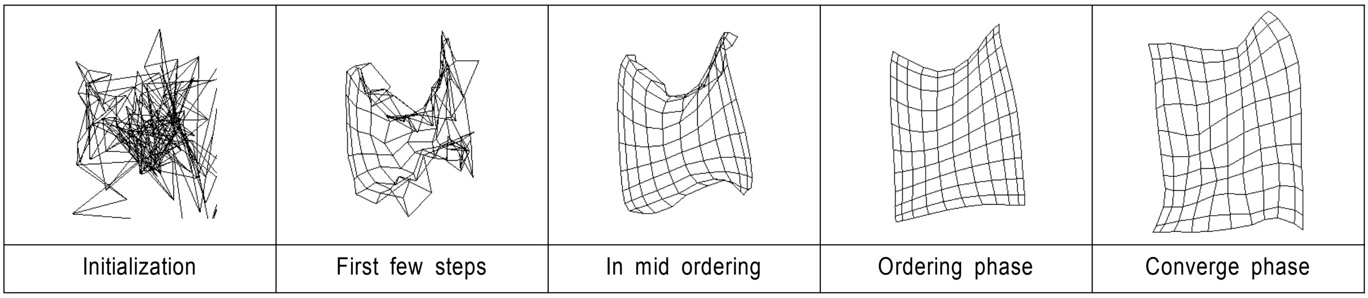
For the performance evaluation, two quantities are considered: (1) we check if the background regions detected contain the foreground objects; and (2) we check if the complimentary areas of background regions do not contain foreground objects. We output 0 if subset of connected component of foreground object is identified as background region and 1 otherwise. In the 23 images, 5348 connected components of foreground objects are found. The connected components are identified automatically using adaptive filtering and connected component algorithm and after that are modified by manually for the regions identified incorrectly. The result is presented using confusion matrix in <Table 3>.

The classification error of identifying foreground as background typically occurs when the input image has very low contrast (the image of a bank check). On the other hand, the classification error of identifying background as foreground occurs when background has a uniform color region locally.

Some of the results of color segmentation studied in this paper are presented. First of all, we showed the reduction of file size when color segmentation is applied. As seen in

〈Table 3〉 The confusion matrix for background detection

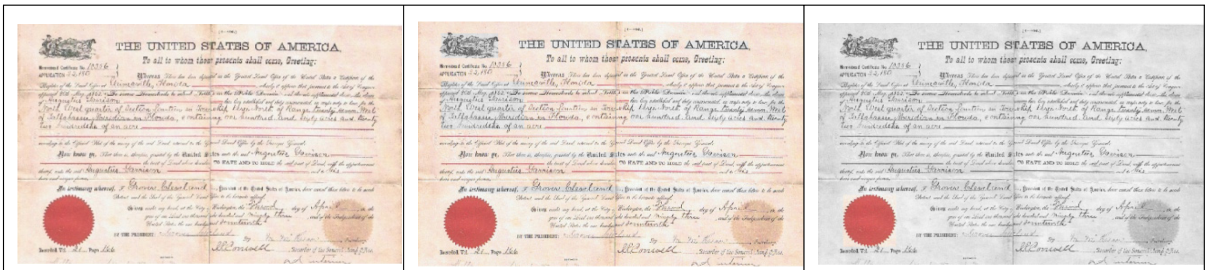
	Background	Foreground
Background (102)	71 (69%)	31 (31%)
Foreground (5246)	158 (3%)	5088 (97%)



(Fig. 2) Visualization of self-organizing feature map construction. For visualization purpose, we only show two-dimension projection of feature map with the first two (R, G) components of input vector.

A. color image	B. color segmentation result

(Fig. 3) The correct result of color segmentation. Color scanned image is saved as gray scale image with color range overlaid.



A. color image	B. conventional method	C. background segmentation

(Fig. 4) The result comparison of color segmentation methods

(Figure 3-A), an input color image has 86Kb in JPEG format. Color segmentation recognizes a region of red-colored stamp. We convert the input image to gray-scale (or bi-tonal) format and then overlay the color recognized region on top of the gray image. The resulting image has 67Kb in JPEG format.

Secondly, we present our segmentation result using the background segmentation. For presentation purpose, we selected an input image having non-uniform color background as seen in (Figure 4-A). The conventional color segmentation methods using pyramid linking, provided by Intel OpenCV library, cannot locate color region effectively. Refer to (Figure 4-B). Background color segmentation shows successful result. Refer to (Figure 4-C).

References

[1] ISO/IEC JTC1/SC29/WG1 N542,"ITU-T Mixed Raster Content model at JPEG2000 Architectural Framework," June 30, 1997.

[2] Gonzalez, R. C. and Woods, R. E., "Digital Image Processing," 2nd Edition, Prentice-Hall, Upper Saddle River, NJ, 2002.

[3] Jähne, B., "Digital Image Processing," Springer-Verlag, Berlin, Germany, 2002.

[4] Lucchese, L. and Mitra, S.K., "Color Image Segmentation: A State-of-the-Art Survey," (invited paper), Image Processing, Vision, and Pattern Recognition, Proc. of the Indian National Science Academy, Vol.67, A, No.2, pp. 207-221, March 2001.

[5] Zhang, Y., "A survey on evaluation methods for image segmentation," Pattern Recognition, Vol.29, No.8, pp.1335-1346, 1996.

[6] Zhang, Y., "A review of recent evaluation methods for image segmentation," Signal Processing and its Applications, 6th International Symposium on 2001, Vol.1, pp. 148-151, 2001.

[7] Hase, H., Yoneda, M., Tokai, S., Kato, J. and Suen, Y., "Color segmentation for text extraction," Int. J. Doc. Anal. Recognit., Vol.6, No.4, pp.271-284, 2003.

[8] Jia, J. and Tang, CK., "Inference of Segmented Color and Texture Description by Tensor Voting," IEEE Trans. Pattern Anal. Mach. Intell., Vol.26, No.6, pp.771-786, 2004.

[9] Lim, J., Park, J. and Medioni, G.G., "Text segmentation

in color images using tensor voting," Image Vision Comput., Vol.25, No.5, pp.671-685, 2007.

[10] Ueda, K., "Extraction of signature and seal imprint from bankchecks by using color information," ICDAR '95: Proceedings of the Third International Conference on Document Analysis and Recognition (Vol.2), p.665, 1995.

[11] Soria-Frisch, A., "Color seal extraction from documents: robustness through soft data fusion," EURASIP J. Appl. Signal Process., No.1, pp.2146-2152, 2005.

[12] Worring, M. and Todoran, L., "Segmentation of Color Documents by Line Oriented Clustering Using Spatial Information," ICDAR, Vol.00, p.67, 1999.

[13] Lucchese, L. and Mitra, S.K., "Unsupervised Segmentation of Color Images Based on k -means Clustering in the Chromaticity Plane," CBAIVL '99: Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries, IEEE Computer Society, Washington, DC, 1999.

[14] Cheng, B., Zheng, N., Wang, Y., Zhang, Y. and Zhang Z., "Color image segmentation based on edge-preservation smoothing and soft C-means clustering," MG\&V, Vol.11, No.2/3, pp.183-194, 2002.

[15] Hammouche, K., Diaf, M. and Postaire, J. -G., "A clustering method based on multidimensional texture analysis," Pattern Recogn., Vol.39, No.7, pp.1265-1277, 2006.

[16] Manjunath, B. S. and Chellappa, R., "Unsupervised Texture Segmentation Using Markov Random Field Models," IEEE Trans. Pattern Anal. Mach. Intell., Vol.13, No.5, pp.478-482, 1991.

[17] Chakraborty, A. and Duncan, J.S., "Game-Theoretic Integration for Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., Vol.21, No.1, pp.12-30, 1999.

[18] Zhu, Q., Li, Y. and He, X., "a Neural Network based color document segmentation approach," Proceedings of the 11th joint international computer conference, pp.925-928, 2005.

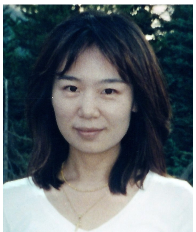
[19] Bruce, J., Balch, T. and Veloso, M. M., "Fast and inexpensive color image segmentation for interactive robots," In Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '00), Vol.3, pp.2061-2066, Oct., 2000.

[20] Hassoun, M. H., "Fundamentals of Artificial Neural Networks," The MIT Press, Cambridge, MA, 1995.

[21] Haykin, S., "Neural Networks - A Comprehensive Foundation," Prentice Hall, Upper Saddle River, NJ, 1999.

[22] Fort, J.C. "SOM's mathematics," Neural Networks, 19: 812-816, 2006.

[23] Kohonen, T., "Self-Organizing Maps", 2nd edition, Springer-Verlag, Berlin, 1997.



신 현 경

e-mail : hyunkyung@kyungwon.ac.kr

2002년 State University of New York (Stony Brook) 대학원 응용수학과 (공학박사)

2008년~현재 경원대학교 수학정보학과 조교수

관심 분야 : Neural network. Machine learning. Image processing