

An Efficient Segmentation Method for Remote Sensing Images Using Self-Organization Map

Jiunn-Lin Wu(吳俊霖) and Shih-Chun Fan-Chiang (范姜士均)
Department of Computer Science, National Chung Hsing University
jlwu@cs.nchu.edu.tw

摘要

本論文提出一個基於類神經網路的彩色遙測之影像分割法。我們提出使用自組織映射圖 (Kohonen self-organizing map) 來萃取影像中主要的特徵，接著利用所得的特徵進行無監督式影像的分類 (unsupervised image segmentation)。在傳統的自組織映射圖，輸入層通常採用像素 (pixel) 本身的數值，並未考慮到影像周遭的像素，所以在森林遙測影像分割上並不能得到滿意的結果。事實上在自然影像中，像素本身與周遭像素都具有很大的相依性，於是此論文中，我們提出一個修正的自組織映射圖演算法，我們在輸入層加入空間特徵 (spatial features)，包含平均值濾波器 (mean-filter)、中值濾波器 (medium-filter) 與離散餘弦變換 (discrete cosine transform) 之係數等。此外我們並且給予每個神經元一個權重值，使其針對不同的輸入，產生對應的權重值，以達到針對不同類型的輸入，均可對應至正確的輸出位置。經過訓練過後的類神經網路，我們提出一有效的後處理的步驟來將具有相同類型的輸出神經元，結合成同一種輸出，並且利用簡單的濾波器來將孤立點移除。我們應用所提演算法來處理森林遙測影像的分割。在實驗中，我們把將不同的樹種視為不同的材質，針對材質的特型，給予不同的特徵，以達到分類的目的。實驗結果顯示所提的方法可以有效的對於彩色影像、遙測影像以及森林影像進行分類。

關鍵字：影像分割，自組織映射圖，類神經網路。

Abstract

In this paper, an efficient segmentation method based on neural network is proposed for the color remote sensing images of forest. It is facilitated by Kohonen self-organizing map (SOM) network, and it performs the unsupervised segmentation.

The images of different of tree species usually have the similar color distribution, and the differences between them are textures. The traditional SOM usually obtains a poor result in the segmentation of forest images, since it uses only the intensities of R, G, and B channels, it does not consider the relationship existed in the neighborhoods of pixels. However, in practice, the pixels in natural images usually have

strong correlation with their neighborhoods. Therefore, we propose a modified self-organizing map network in this paper, it uses the additional spatial features in the input layer, such as the coefficients of discrete cosine transform. In this way, we consider both pixels themselves and the correlation information with their neighborhoods at the same time. We also add a new weighting function for each neuron, which can help each neuron to map to a suitable output neuron. Finally, we use the noise-filter to improve segmentation quality at the post-processing stage. Experimental results show that the proposed method can separate successfully the different color texture in the remote sensing images of forest.

Keywords : Image segmentation, SOM, DCT, Remote sensing, Forest.

1. Introduction

Digital aerial photography can be used to monitor growth, fire problems, insecticide coverage, fertilization, and irrigation of forest lands. The image segmentation is a very important technique when analyzing the remote sensing image of forest. Image segmentation is a process that divides different regions or objects of the image, and thus we can retrieve the region of interest. Image segmentation partitions the image into different meaningful regions with homogeneous characteristics.

According to the usage of prior knowledge of the image, color image can be segmented in the unsupervised or supervised way [3]. The unsupervised segmentation is usually used in the applications that the image features are unknown, for example, remote sensing image analysis, nature scene understanding etc. The supervised segmentation is commonly used in the applications that we have sample of object color features, such as object tracking, face recognition, and image retrieval [6]. In this paper, we propose an efficient segmentation method for the color remote sensing image of forest, it is based on Kohonen self-organizing map (SOM) network, and it performs the unsupervised segmentation. The motivation for this architecture arose from the need to develop a strategy for color image segmentation for which the amount of training data may be limited.

The traditional self-organizing map networks

usually use the intensities of the pixels as features in the input layer [4][9]. However, in practice, they get poor segmentation results for the remote sensing image of forest. The images of nature senses usually have strong correlation between pixel neighborhoods. Adjacent pixels in an object generally depend on each other. For overcoming the weakness of the traditional SOM, we will use the spatial features, such as the mean and variance of a neighborhood, the output of the median filter, the coefficients of DCT and so on. Experimental results show that the proposed method performs successfully on the segmentation of the color remote sensing images of forest.

The rest of this paper is organized as follows: We describe the fundamental concepts of the image segmentation approaches based on the neural network in Section 2. In Section 3, the structure of the modified SOM is proposed. Experimental results and discussion are given in Section 4. Finally, the conclusion is given in Section 5.

2. Background

The Neural network is a mathematical or computational model. It used large numbers and simple artificial neuron to simulate human neural. The neuron gets information from input or other neurons, then uses this information to compute result, and passes this result to other neurons or output. The Kohonen self-organizing map (SOM) is an unsupervised learning method [1]. In image segmentation, we can use SOM to capture the dominant colors of the image. Since SOM is an unsupervised segmentation method, we don't need any priori knowledge for the test images [4].

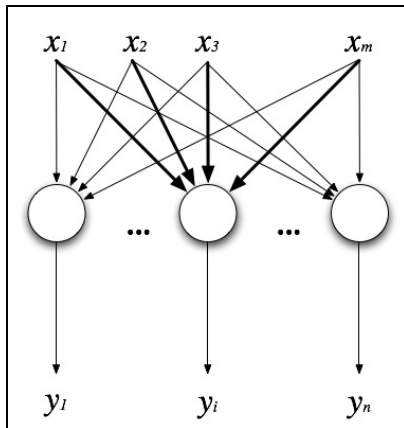


Fig. 1. The Self organization map network.

The self-organizing map is structured as a two-layer neural network as shown in Fig.1. It consists of one input layer, and one output layer. Every input neuron connects with every output neurons. The dimension of input layer is according to input features. The dimension of input layer is according to input features. The self-organizing feature map network use

the Kohonen (winner-take-all) learning rule. Winner-take-all is the learning rule that the network will automatic chooses the winner during each iteration. When the winner is chosen, the network will update all the weight value that connects with the winner. The SOM can be applied to unsupervised image segmentation. However, the conventional SOM usually use the intensities of R, G and B in color images as the input feature, they get poor segmentation result when analyzing the aerial images of forest. It is because of the images of different of tree species usually have very similar color distribution, and the only difference between them are the texture maps.

The discrete cosine transform (DCT) is a Fourier-related transform, it is similar to the discrete Fourier transform (DFT), but using only real numbers [2]. It is equivalent to a DFT of roughly twice the length, operating on real data with even symmetric, where in some variants the input and/or output data are shifted by half a sample. Since the DCTs are orthogonal transform representations, they have properties similar in form to those of the DFT. In this paper, we propose a modified SOM network to improve the drawback of the conventional SOM. We use not only the intensities of R, G and B but also the DCT coefficients in neighborhoods in which they represent the frequency characteristics of the forest image of different tree species.

3. The Proposed Method

It is well known that all colors are perceived as combinations of three primary colors: red, green and blue (RGB). Almost all of today's color image acquisition devices, for example sensors for remote sensing images, output images in RGB space. However, RGB space is not always ideal for different segmentation algorithm since they have strong correlation. In this paper, we adopt the $L^*u^*v^*$ color space where L^* represents the intensity, and u^*v^* represents the color chromaticity. The raw data of the input color image can be converted to $L^* u^*v^*$ color space by the following CIE standard formula [7]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The $L^* u^*v^*$ are given by

$$L^* = \begin{cases} 903.3 \left(\frac{Y}{Y_0} \right) & \text{if } \frac{Y}{Y_0} < 0.008856 \\ 25 \left(100 \frac{Y}{Y_0} \right) - 16 & \text{if } \frac{Y}{Y_0} \geq 0.008856 \end{cases} \quad (2)$$

$$u^* = 13L^* (u' - u'_0)$$

$$v^* = 13L^* (v - v'_0)$$

where

$$\begin{aligned} u' &= \frac{4X}{X+15Y+3Z} \\ v' &= \frac{9Y}{X+15Y+3Z} \end{aligned} \quad (3)$$

and $Y_0 = 225$, $u'_0 = 0.0200953$ $v'_0 = 0.460900$.

Similar with the conventional SOM, the proposed modified SOM has two modes of operations: training process and mapping (classification) process. In the traditional SOM, both two processes are needed to choose as the winner neurons. In the training process, the winner should be chosen for updating weight. In the mapping process, the winner should be chosen in order to replace the input vector. The winner neuron usually is calculated by Euclidean distance. The network calculates the Euclidean distance of each neuron, and then chooses the minimal distance of all neurons. As mentioned above, the pixel intensities have strong correlation with the one of the neighbor pixels in the nature images. In the proposed modified SOM algorithm, we will add spatial features in the input layer. We give each neuron an additional weight as shown in (4).

$$\begin{aligned} \mathbf{w}^{(k+1)} &= \mathbf{w}^{(k)} + \alpha(\mathbf{D}) \\ \mathbf{D} &= [\beta_1(x_1 - w_1) + \beta_2(x_2 - w_2) + \dots + \beta_n(x_n - w_n)] \end{aligned} \quad (4)$$

where w is the weight matrix, α is learning rate, and β is the additional weight we added. For example, if we use three additional features: the mean and the DCT coefficients f_{01} and f_{10} of a neighborhood, the D can be obtained by

$$\begin{aligned} D &= [\beta_1(r - w_1) + \beta_2(g - w_2) + \beta_3(b - w_3) + \\ &\beta_4(\text{mean} - w_4) + \beta_5(f_{01} - w_5) + \beta_6(f_{10} - w_6)] \end{aligned} \quad (5)$$

where r , g and b are the intensities of R, G and B channels respectively. The weight β will be adjusted by different input vectors, and it plays an important role in our algorithm. Assume we use a window whose size is n -by- n , the variance of pixel $p(i, j)$ is calculated by:

$$\text{var}(i, j) = \frac{1}{n \times n} \sum_{k=-(n/2)}^{n/2} \sum_{l=-(n/2)}^{n/2} (x(i+k, j+l) - \bar{x})^2 \quad (6)$$

If the variance $\text{var}(i, j)$ is small, it means the pixel $p(i, j)$ should in a smooth region. On the contrary, if the value of $\text{var}(i, j)$ is large, $p(i, j)$ must be on the edges or in the textures. Obviously, the value of the weight β will depend on the $\text{var}(i, j)$. If $p(i, j)$ is in the smooth region, the β_i corresponding to the pixel value would have the larger

value. Conversely, the value of β_i corresponding to the spatial features will have the larger value if $p(i, j)$ locates on the edges or in the textures.

The proposed SOM uses the Euclidean distance to choose the winner and update the weight matrix. The steps include:

Step.1 Initialization: Choosing some values for the initial weight vectors $\mathbf{w}_j^{(0)}$ where $j=1,2,\dots,N$; where N is the number of neurons in the input layer. The learning-rate $\alpha(0)$ should be initialized with values close to unity. And the additional weight must satisfy

$$\sum_i \beta_i = 1 \quad (7)$$

And the neighborhood size N_i should be half of network size. i is the time of each iteration.

$$\begin{aligned} N_{i_1} &\supset N_{i_2} \supset N_{i_3} \supset N_{i_4} \supset \dots \\ \text{where } i_1 &< i_2 < i_3 < \dots \end{aligned} \quad (8)$$

Step.2 Choose winner: In image segmentation, the output neurons with weights $w_j^{(t)} = [w_{j1}, w_{j2}, \dots, w_{jm}]$ compete with each other to find a best match with the input pattern. A widely used measure for the match of x with w_i is Euclidean distance between them as shown in (9).

$$\|\mathbf{x} - \mathbf{w}_i\| = \{\beta_1(x_1 - w_{i1})^2 + \beta_2(x_2 - w_{i2})^2 + \dots + \beta_m(x_m - w_{im})^2\}^{1/2} \quad (9)$$

The neuron C whose weights vector w_C is closest to x is declared the winner, i.e.

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\} \quad (10)$$

Step.3 Weight update: The winner and all neurons within $N_c^{(k)}$ are updated by the learn rule:

$$\mathbf{w}_i^{(k+1)} = \begin{cases} \mathbf{w}_i^{(k)} + \alpha(k)[\beta_1(x_1 - w_{i1}) + \beta_2(x_2 - w_{i2}) + \dots + \beta_m(x_m - w_{im})] & \text{if } i \in N_c^k \\ \mathbf{w}_i^{(k)} & \text{if } i \notin N_c^k \end{cases} \quad (11)$$

where $\alpha(k)$ is the respective learning rate of the winning neuron and its neighbors at the k th iteration, and $0 < \alpha(k) < 1$. The learning rate $\alpha(k)$ is linearly decreased to $\alpha(t+1) = \alpha(t)(1.0 - t/T)$.

Step.4 Convergence criterion: Learning is a stochastic process, and the final accuracy of the mapping depends on the number of steps. In our scheme, we set a threshold θ . The training process is completed when $\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)}$ is less than θ .

Fig. 2 shows the flowchart of the proposed SOM

network. We add spatial features in the input vector of SOM, for example, we add the mean and variance as the spatial features for the segmentation of the usual color images.

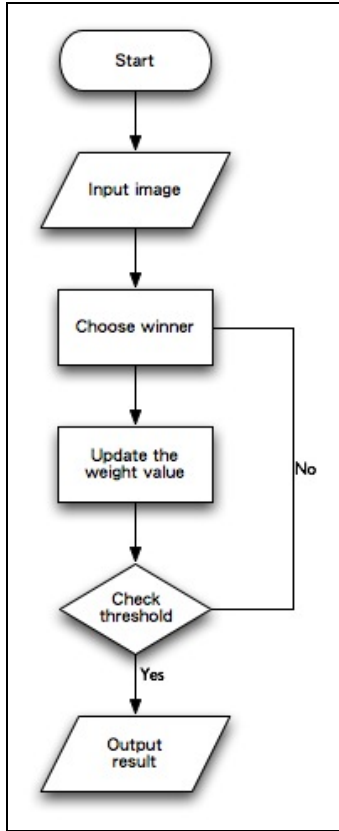


Fig. 2. The flowchart of the proposed SOM.

For the remote sensing images, our focus is on the aerial images of forest. Our target is to separate the different species of trees. As mentioned above, the images of different of tree species usually have the similar color distribution, and the differences between them are textures. So we propose to add the DCT coefficients in the input layer of SOM. The coefficients of DCT are represented the different frequency. Equation (12) is the formula of 2D DCT transform, where the first coefficient represents f_{00} , the second coefficient represents f_{01} , ...,etc. Here, f_{00} represent the mean.

$$f_{i,j} = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{(2x+1)j\pi}{2N}\right] \cos\left[\frac{(2y+1)i\pi}{2N}\right] \quad (12)$$

In this paper, we use the coefficient f_{01} and f_{10} for the segmentation of forest remote sensing images, which are the first two major coefficients in frequency domain of DCT transform. The coefficients have the different values in different texture images.

Fig. 3 shows the segmentation result of an artificial texture images by the proposed method, in which two spatial features, the DCT coefficients f_{01} and f_{10} , are used. Fig. 1(a) is the input image, it

combined with two textures, in which the right part is the vertical texture, and the left is the horizontal texture. Fig. 1(b) is the segmentation result. Obviously, although these two texture maps in Fig. 3(a) have the same colors, the proposed modified SOM can separate them successfully. It demonstrates the additional spatial features, the coefficients of DCT, play an important role in the segmentation of two texture images with same color

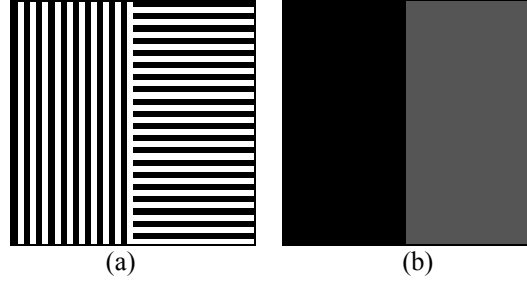
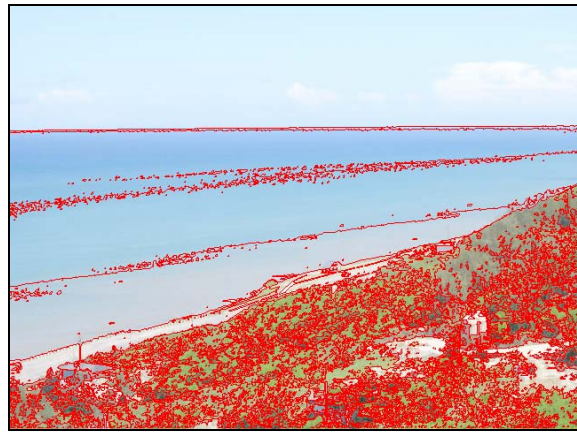
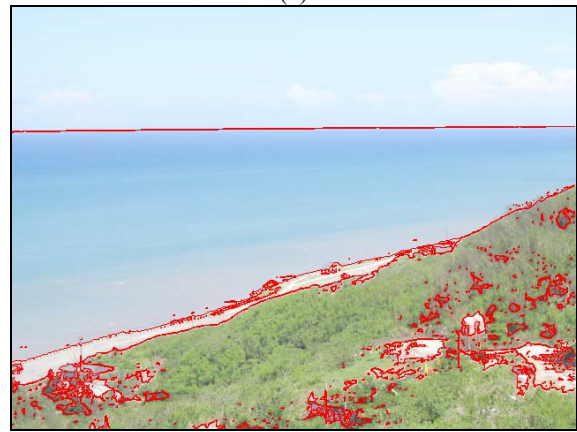


Fig. 3. (a) An image with two different textures. (b) The segmentation result by the proposed SOM using two DCT coefficients.



(a)



(b)

Fig. 4. The segmentation result after the proposed merge step.

We found the results by the proposed method are often over-segmented, there are too many small regions as shown in Fig. 4(a). To solve this problem, we propose a simple merge method. After the training process of the proposed SOM, we calculate the

Euclidean distance between each neuron by (13).

$$\text{distance}(i,j) = \left\| (w_{i1} - w_{j1}), (w_{i2} - w_{j2}), (w_{i3} - w_{j3}) \right\| \quad (13)$$

If the distance is smaller than the predefined threshold, the algorithm will map these two neurons into the same output neuron. After merging neurons, we use a simple noise reduction filter to remove the isolated points. Fig. 4(b) is the segmentation result after the merge step. It demonstrates the proposed merge step is able to get a relatively rough segmentation result, and every rough region is composed of several detailed regions.

4. Experimental Results

To verify the effectiveness and robustness of our proposed method, we used some color remote sensing image of forest for testing in the experiments. The experiments were conducted using Mathworks Matlab 7.0 programming on a Pentium 4 CPU 3.0GHz platform with 512MB memory. In the experiments, we use the proposed SOM network to segment the texture images with different plants or trees. The experimental aerial images are from ‘‘Aerial Survey Office, Forestry Bureau, and Council of Agriculture Executive Yuan’’ [9].

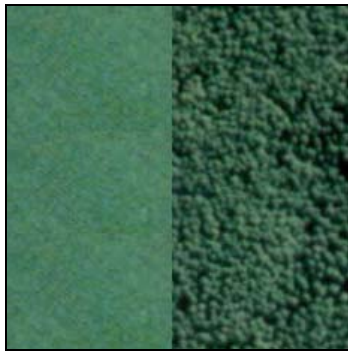


Fig. 5. A texture image of different plant species.

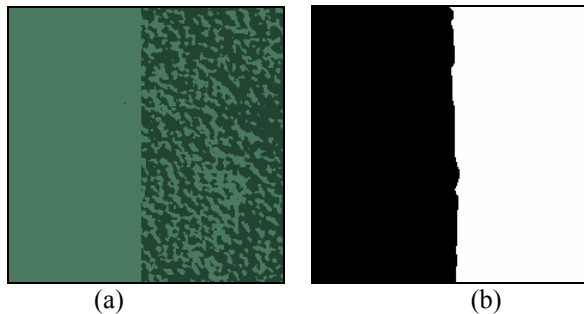


Fig. 6. The segmentation result of Fig.5. (a) The result when the input features are the intensities of RGB. (b) The result with three additional DCT coefficients as input features.

Fig. 5 shows an aerial image with two different

plants. The left side texture is the *corn*, and the right side is the *cryptomeria*. The image size is 230×230 pixels. If we only use the intensities of R, G and B as the input features, i.e. the input feature is $\mathbf{x} = [x_1, x_2, x_3]$, the segmentation result is shown in Fig.6(a). Obviously, the left part of the image is segmented clearly, but the right part is failed. It means if we use only the pixel intensities, the segmentation algorithm can't separate the image correctly. For overcoming this problem, we use three additional DCT coefficients as the input features. Now the dimension of the input features is 6, i.e. $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5, x_6]$ where the first three features are pixel values of R, G and B, and the last three features are the coefficients of DCT. The segmentation result is shown in Fig.6(b). Obviously, the proposed method performs good segmentation results, the textures of different plants are segmented clearly. It demonstrates that the DCT coefficients do help us to segment the texture images successfully.

Fig.7 is another plant texture image, in which the left side texture is *corn*, and the right side texture is *Eucalyptus*. The image size is 230×230 pixels. Fig. 8(a) is the result if we use only the intensities of RGB, and Fig.7(b) shows the segmentation result when using three additional DCT coefficients as input features. Obviously, the proposed method gets a satisfied segmentation result now.

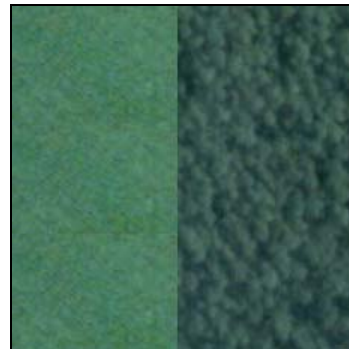


Fig. 7. Another texture image of different plant species.

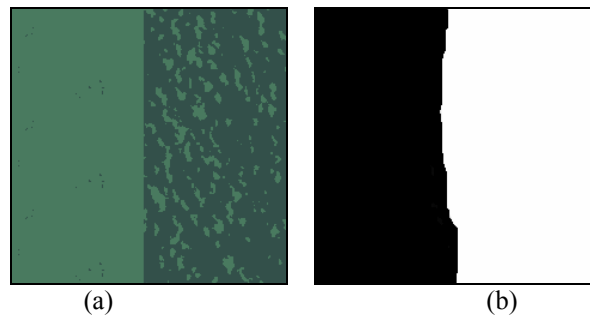


Fig. 8. The segmentation result of Fig. 7. (a) The result when the input features are the intensities of RGB. (b) The result with three additional DCT coefficients as input features.

Fig. 9 shows an aerial remote sensing image of forest from [9], it contains at least five different tree or plant species. The image size is 353×313 . Fig. 10 is the segmentation result by the proposed method, in which we use three DCT coefficients as the input features of SOM. Obviously, the different tree species are separated successfully. It demonstrates the effectiveness and robustness of the proposed modified SOM. The computation time using our modified SOM is around 1-2 minutes, it is much faster than the typical segmentation task such as FCM method for the color images.



Fig. 9. A remote sensing image of forest.

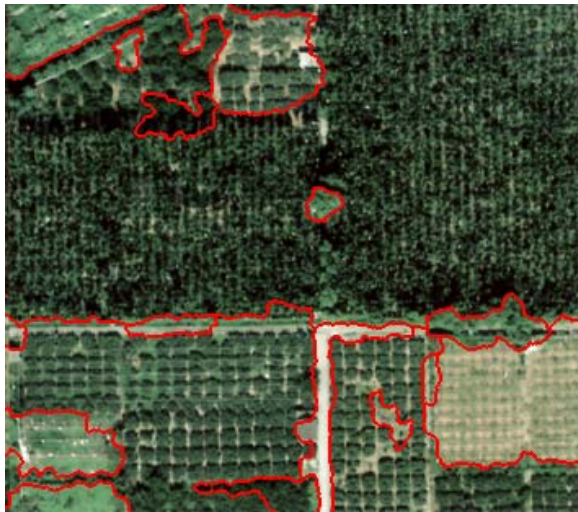


Fig. 10. The segmentation result of Fig. 9 by the proposed modified SOM.

5. Conclusion

In this paper, we propose an unsupervised segmentation system based on self-organization map for the color remote sensing images. It uses the spatial features to improve the segmentation result. Since the DCT coefficient represents the frequency

characteristic of the texture images, we use the additional the DCT coefficients as the input features of SOM network. It helps the network performs better than the conventional SOM, especially in the case of objects with different texture maps but with same color. Finally, we propose a post-processing scheme based on the Euclidean distance of neurons to yield a rough segmentation result. If the distance is smaller than the predefined threshold, the algorithm will map these two neurons into the same output neuron. After merging neurons, we use a simple noise removal filter to remove the isolated points.

Experimental results show that the proposed method can segment the color remote sensing image of forest correctly. In comparison with other segmentation algorithm, the processing time of the proposed method is also fast.

Reference

- [1] G. Dong and M. Xie, "Color Clustering and Learning for Image Segmentation Based on Neural Network" *IEEE Transactions on Neural Networks*, vol.16, no. 4, pp. 925-936, 2005.
- [2] T. Kohonen, Ed., "Self-Organizing Maps", *Berlin, Germany: Springer -Verlag*, 1995.
- [3] C. T. Lin, C.S. George Lee, *Neural Fuzzy System*, Prentice Hall, New Jersey, 1996.
- [4] S.H. Ong, N.C. Yeo, K.H. Lee, Y.V. Venkatesh, D.M. Cao "Segmentation of color images using a two-stage self-organizing network" *Image and Vision Computing*, vol. 20, pp. 279-289, 2002.
- [5] A.V. Oppenheim, R.W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing, second edition*. Prentice-Hall, New Jersey, 1999.
- [6] N.R. Pal and S.K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, pp. 1277-1294, 1993.
- [7] A.R. Robertson, "The CIE 1976 color-difference formulae", *Color Research and Applications* 2, pp. 7-11, 1977.
- [8] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map" *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 3, pp. 586-600, 2000.
- [9] "航照立體像片對" 行政院農業委員會林務局農林航空測量所印行 叢刊第 108 號 中華民國 94 年 11 月