ISSN No. 0976-5697

Volume 5, No. 8, Nov-Dec 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Automatic Segmentation of Satellite Image using Self Organizing Feature Map (SOFM)An Artificial Neural Network (ANN) Approach

Anil Kumar Goswami
DRDO
Delhi, India

Rashi Jain
Banasthali University
Jaipur, Rajasthan, India

Priyadarshni Tripathi
Banasthali University
Jaipur, Rajasthan, India

Abstract: Satellite Image Processing is one of the key areasof computer science which is used for various important applications of remote sensing and Geographic Information System (GIS). Image segmentation is one of the key research areas in the field of image processing and analysis. Image segmentation is needed for various intermediate as well as final processes for various analyses. There are various algorithms available for image segmentation. Self Organizing Feature Map (SOFM), an Artificial Neural Network, is one of such algorithms. Based on the dataset, SOFM performs the training of the neural network and then segment the input satellite image into various homogenous clusters. Later the similar regions are merged and results are refined. This paper presents detailed working of SOFM together with the experiment conducted for satellite image segmentation.

Keywords:Artificial Neural Network, Self Organizing Feature Map (SOFM), Image Processing, Segmentation, Clustering

I. INTRODUCTION

Image processing and analysis is one of the key areas of computer science research and applications. One important aspect of image processing and analysis is segmentation of satellite images. Segmentation is useful for extracting of important and relevant information from data. It can be defined as clustering of data into several homogenous regions based on certain criteria. It groups the data in several regions on the basis of similarities in some features of data. The features could be of various types such as RGB values in images or some attributes calculated using RGB values. There are various algorithms for image segmentation. These algorithms may be grouped into many categories such as region-splitting and merging based segmentation, data clustering, edge based segmentation, and threshold based segmentation[1].

SOFM is an algorithm for image segmentation, developed by Kohonen in 1990. It is a type of Artificial Neural Network which exploits the learning capability of neurons. Learning is of two types namely supervised and unsupervised. SOFM make use of unsupervised learning and generates memory of the network which is then used to do the segmentation of satellite image. In unsupervised learning, there is no prior knowledge except data input, hence it is necessary to have network that learns self and generates accurate result by extracting the information from input data.

This paper presents image segmentation using SOFM. The second section of paper presents literature review whereas the third section presents basic understanding of SOFM. Methodological framework including flow charts and algorithm has been explained in fourth section of the paper. Fifth section presents results and discussion whereas conclusion is given in last section.

II. LITERATURE REVIEW

An image may be considered to contain sub images sometimes referred to as ‘regions’. This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region [2]. The clustering process aims to find the underlying structures and partition the collection into clusters of ‘similar’ images [3]. Image segmentation is the process of division of the image into regions with similar attributes [4].The image segmentation can be categorized into (1) edge and line oriented segmentation (2) region growing method (3) clustering and (4) region splitting methods [1].Applications of image segmentation include Content-Based Image Retrieval (CBIR), object recognition, matching of stereo pairs for 3-D reconstruction, navigation, artificial expert medical diagnosis etc. [5]. Self-organizing neural networks are designed to detect regularities and correlations in the input they receive and adapt their future responses to that input accordingly [6]. In the area of bioinformatics, SOFM has been used primarily for visualizing protein and DNA sequence and structure spaces [7].

SOFM concept was introduced as a feature-extraction and data-mapping approach by Kohonen in 1982 [8].SOFM can be used as an effective tool for data mining, in particular for clustering algorithm [9]. SOFM too has some disadvantages. If adjacent map units point to adjacent input data vector, so sometimes distortions are possible because high dimensional topography cannot always be represented in 2D. To avoid such phenomenon, training rate and the neighborhood radius should not be reduced too quickly [10].

III. SELF ORGANIZING FEATURE MAP (SOFM)

SOFM is an Artificial Neural Network known as Kohonen network. It is named as “Self-Organizing” because no supervision is required for its training. It learns on its own by unsupervised learning. The input data to SOFM can be of any dimensions. It consists of network of nodes known as neurons. Each node in network contains a corresponding weight vector.There are generally three phases involved in the formation of SOFM which are 1) competition, 2) co-operation and 3) adaption[11]. The input features for SOFM are represented by the chromaticity values, as it specifies quality of a color regardless of its luminance. Chromaticity is obtained by normalizing the band for input image. Chromaticity is computed using equation 1.

$$Chromaticity\ of\ i_{th}\ band = \frac{value\ of\ i_{th} band}{Sum\ of\ values\ of\ all\ bands} \quad (1)$$

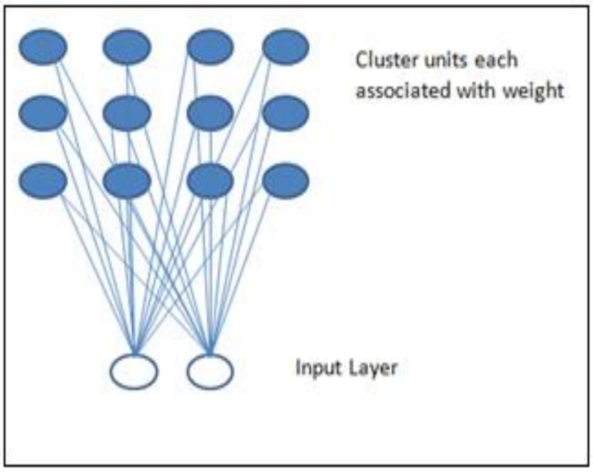

Figure 1. SOFM Network Architecture

During the training phase, all the nodes compete for the selection process of the Best Matching Unit (BMU). The input vector *x* at time *t* is compared to each of the weight vectors w_i and the minimum euclidean distance between the inputs and the nodes weight determines the closest match.

$$\text{Euclidean Distance} = \|x(t) - w_i(t)\| (2)$$

When the BMU is determined, the nodes closest to BMU tend to get excited more than those which are far away. There is a scalar function gain function Θ(t) that decays with distance [12].

$$\theta(t) = e^{\frac{(-dist)^2}{2\times(neighborhood\ radius)^2}} \quad (3)$$

Where
dist=distance of input vector to BMU

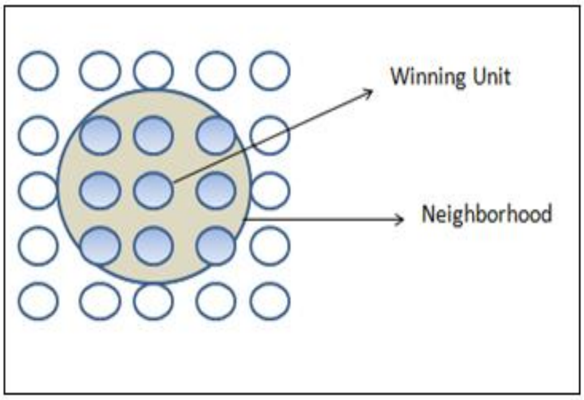

Figure 2. SOFM Neighborhood around a winning unit

Neighborhood radius is calculated using formula given in equation 4.

$$Neighborhood\ Radius = \text{Clusters} \times e^{\frac{-\text{Current epoch}}{\text{Time constant}}} \ (4)$$

where

$$Time\ Constant = \frac{Total\ no.of\ epoch}{log(Clusters)} \ (5)$$

Here time constant is a constant unit while neighborhood radius will decay corresponding to increase in time as time will increase current no of epoch will increase so the negative of exponential function will decrease as result of which radius will also decrease.

This decay is due to the fact that neighborhood radius decays over time. This gain function is maximum at the winning nodes and decreases with the distance. Outside the neighborhood radius, the nodes are not affected.

In this experiment, the neighborhood concept has been applied as shown in figure below.

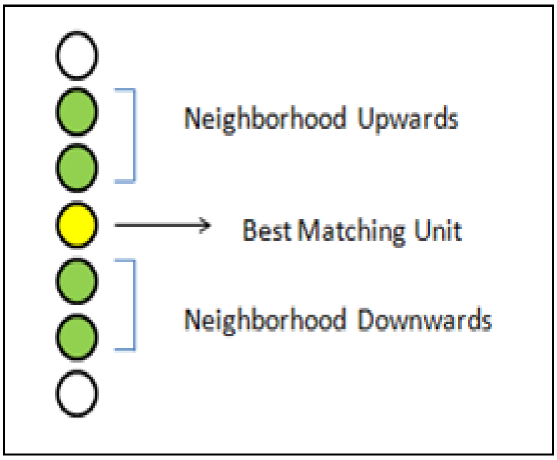

Figure 3. SOFM Neighborhood applied

During the adaption (learning phase) the node weights are updated. Adjacent nodes of winner nodes respond similarly, while distant nodes respond diversely [12]. The weights are updated considering a neighborhood around the BMU. Scalar function gain function θ(t) is employed to establish how the updating will be done inside the neighborhood. L (t) is the learning factor of the network. The weights are updated using equation 6.

$$w_i(t+1) = w_i(t) + \theta(t)L(t)[x_i(t) - w_i(t)] (6)$$

The process of training and adaption keeps going on until optimal solution is achieved. Once optimal solution is achieved, it results into number of homogenous regions. Then, the process of region merging takes place in which similar regions are merged based on threshold value computed using equation 7.

$$ThresholdRegionMerging = MeanOfDistanceBetween\ MeanOfClusters - StandardDeviationOfDistance\ BetweenMeanOfClusters (7)$$

where standard deviation (σ) is

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} (8)$$

xi= mean of Bands Of Clusters
μ=MeanOfDistanceBetweenMeanOfClusters
N= TotalNoOfDistanceBetweenMeanOfClusters

$$N = \frac{clusters \times (clusters - 1)}{2} \quad (9)$$

If the distance between means of regions is smaller than the threshold value, regions are merged. This is necessary for getting rid of over segmentation.

IV. METHODOLOGICAL FRAMEWORK

The basic objective of the proposed work is to do the segmentation of satellite image intovarious clusters. A satellite image is a matrix of pixels where each pixel is having multiple values depending upon the number of bands in it. It is given as input and the weights of the nodes of network layer are initialized to small random values.

The training of SOFM occurs and the weights of nodes are updated. It includes the calculation of the chromaticity values of the images and then finding the BMU. Then the amount of influence is calculated and accordingly the weights of the nodes are updated. This process continues till the optimum solution is achieved.

Then, the merging of the clusters is done based on the threshold value. Threshold value is calculated with the mean and the standard deviation of the pixel's values of various clusters. If the value of a cluster is less than the threshold value, the clusters are merged.

In the end, the complete segmented image is obtained and color codes corresponding to each cluster are assigned.

A. *Block Diagram*

The block diagram of the system is given below. There are four important components of SOFM namely Best Matching Unit Computation, Clusters weight updation, Initial region assigned and Region Merging.

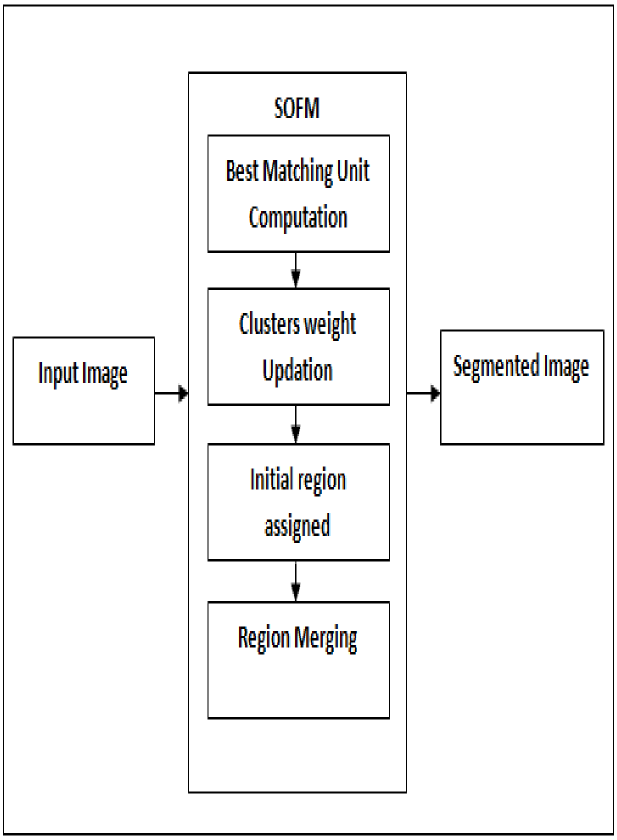

Figure 4. Block Diagram

B. *Flow Chart*

The flow chart of SOFM is given below.

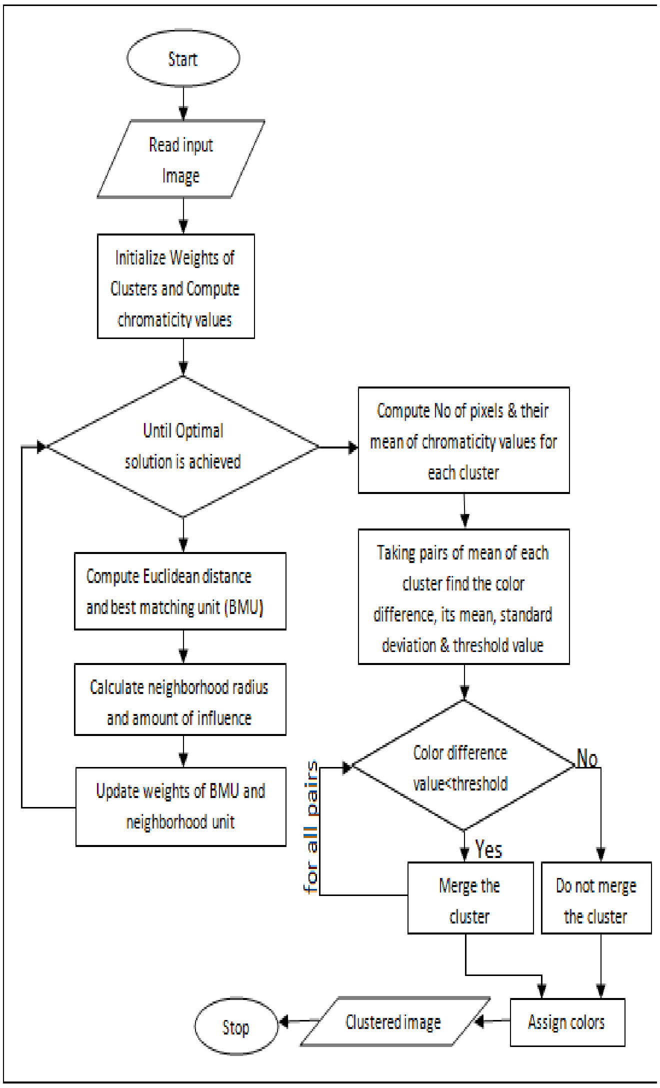

Figure 5. Flow Chart

C. *Algorithm*

The algorithm of SOFM is given below.

1) **Inputs**
   Image Height, Image Width, NoOfBands, NoOfClusters, Learning Factor, DataMatrix[ImageHeight][ImageWidth][NoOfBands]

2) **Initialize the weights of the Network layer.**

3) **Repeat until optimal solution is achieved.**
   3.1) Compute Chromaticity values

$$Chromaticity\ of\ i_{th}\ band = \frac{value\ of\ i_{th} band}{Sum\ of\ values\ of\ all\ bands}$$

   3.2) Calculate the Euclidean Distance between input vector x and the weight vectors w_i

$$Euclidean\ distance = \|x(t) - w_i(t)\|$$

   3.3) Find the Best Matching Unit(BMU)

$$BMU = min\|x(t) - w_i(t)\|$$

   3.4) Determine the Neighborhood radius

$$Neighborhood\ Radius = NoOfClusters \times e^{\frac{-Current\ epoch}{Time\ constant}}$$

where

$$Time\ Constant = \frac{Total\ no.of\ epoch}{log(Clusters)}$$

   3.5) Calculate the amount of influence θ(t)

$$\theta(t) = e^{\frac{(-dist)^2}{2\times(neighborhood\ radius)^2}}$$

where
dist = distance of input vector to BMU

   3.6) Update the weights of BMU and its neighborhood

$$w_i(t+1) = w_i(t) + \theta(t)L(t)[x_i(t) - w_i(t)]$$

where
θ(t)=AmountOfInfluence
L(t)=LearningFactor

4) **Calculate Chromaticity values, compute BMU and find initial regions.**

5) **Find the mean of each region.**

$$MeanOfBandsIfClusters = \frac{SumOfBandsOfClusters}{NoOfPixels\ for\ a\ cluster}$$

6) **Taking pairs of regions, compute the total Number of Distancesbetweenmean ofregions.**

$$TotalNoOfDistanceBetweenMeanOfClusters = \frac{clusters\times(clusters-1)}{2}$$

7) **Calculate the mean of all distances between means of regions.**

$$MeanOfDistanceBetweenMeanOfClusters = \frac{SumOfDistanceBetweenMeanOfCluster}{TotalNoOfDistanceBetweenMeanOfCluster}$$

8) **Calculate the Standard Deviation of all the Distance between mean of regions.**

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

where
x_i=mean of Bands Of Clusters
μ=MeanOfDistanceBetweenMeanOfClusters
N=TotalNoOfDistanceBetweenMeanOfClusters

9) **Calculate Threshold value for region merging.**

$$ThresholdRegionMerging = MeanOfDistanceBetween\ Mean\ OfClusters - StandardDeviationOf\ Distance\ BetweenMeanOfClusters$$

10) **Merge those regions where distance Between Mean of regions is smaller than the Threshold Value.**

11) **Assign color codes to each region.**

11) **Output**- Segmented Image

V. RESULTS AND DISCUSSION

The results of above experiment have been shown in this section.

A. *Data Distribution*

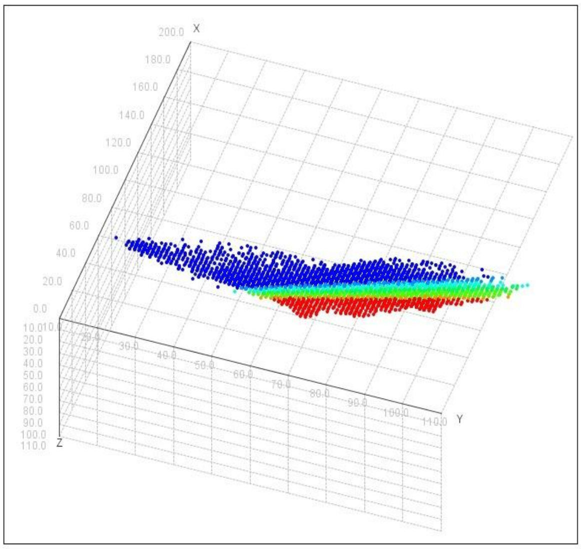

Figure 6. Data distribution of each regions

Figure 7 shows the data distribution of each regions along three dimensions namely R, G and B representing three values of each input pixel. It clearly shows the grouping of similar data into a region, although some overlapping may exist depending on the input data distribution.

B. *Change in neighborhood radius*

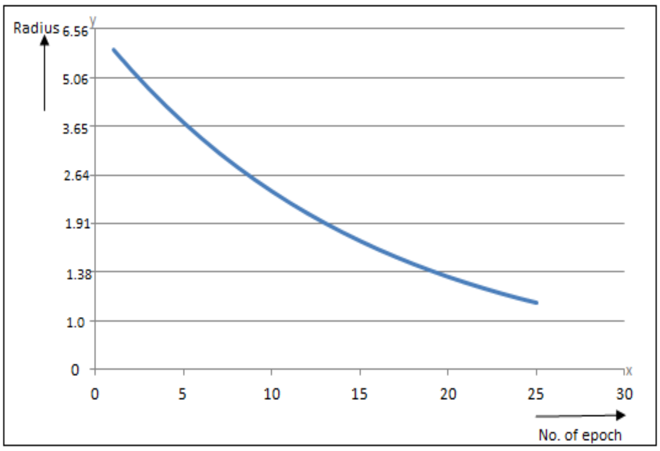

Figure 7. Change in Neighbourhood Radius

C. *Segmented Images*

Input images and resultant segmented images are shown in figure 8. The color codes of segmented images represents the homogenous regions.

The results of SOFM have shown the strength of SOFM as a good algorithm for image segmentation. It is a powerful algorithm and being used for various applications.

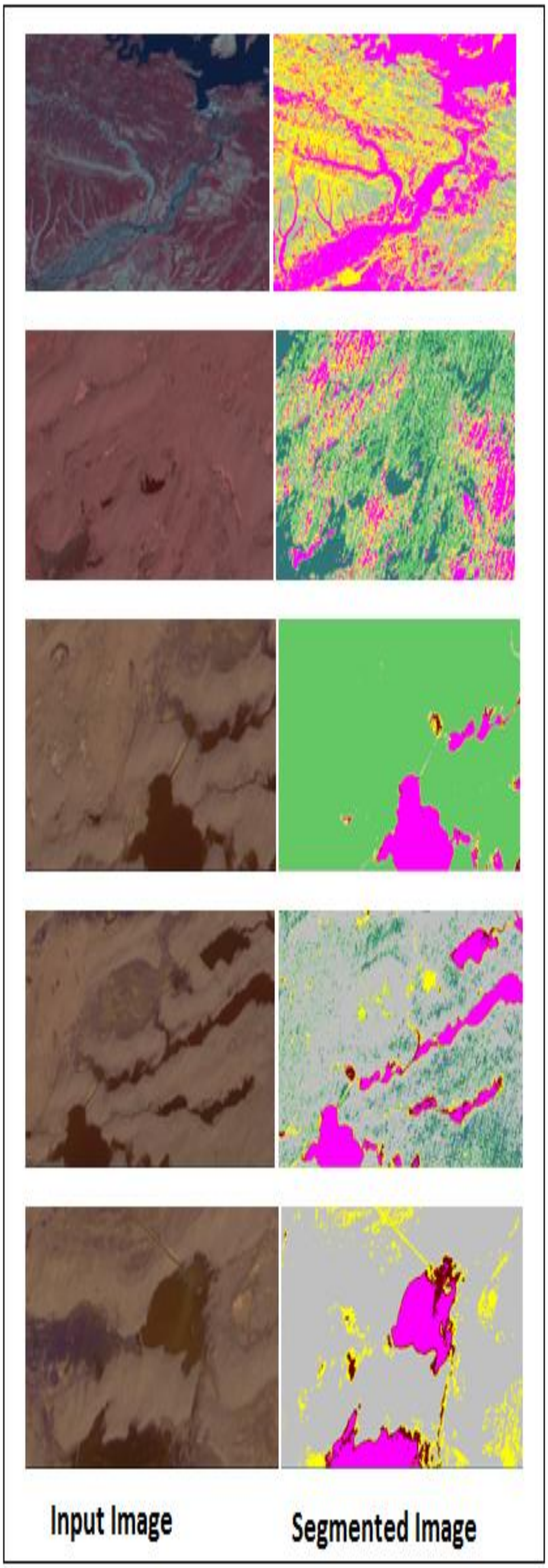

Figure 8. Segmented Images.

VI. CONCLUSION

There are various image segmentation algorithms. One of such algorithms is SOFM. This study has used SOFM, an unsupervised image segmentation algorithm, for satellite image segmentation. SOFM has been implemented and tested with satellite images. The results obtained show the strength of SOFM as a good algorithm for image segmentation. The beauty of SOFM lies in its training and region merging unlike other algorithm such as K-Means. Region merging step of SOFM reduces total number of regions and hence resulting into significant regions only. For various applications, the first step is generally image segmentation and SOFM can be very useful for this purpose. Further, SOFM can be used in combination with other algorithms to get more consistent and accurate results.

VII. REFERENCES

[1] Y. I. Ohta, T. Kanade and T. Sakai, “Color information for region segmentation”, Computer graphics and image processing,vol. 13,pp 222-241, 1980.

[2] Lan T. Young, Jan J. Gerbrands and Lucas J. Van Vliet,,“Fundamentals of image processing”,1st edition, Version 2.3,pp 1-112,1995-2007.

[3] Alon Faktor and Michal Irani, “Clustering by composition”, Springer-Verlag Berlin Heidelberg, pp 474-487, 2012.

[4] William K.Pratt., “Digital image processing”, 2nd edition. John Wiley & Sons, 1991.

[5] M. Arfan Jaffar, Muhammad Ishtiaq, Ayyaz Hussain and Anwar M. Mirza, “Wavelet-Based Color Image Segmentation using Self-Organizing Map Neural Network*”,*International Conference on Computer Engineering and Applications, vol.2, pp 430-434, 2009.

[6] Karishma Tyagi and Gaurav Sachan, “A comparison of SOFM with K-Mean and ANN Pattern Reorganization”, International Journal of Engineering Research and Development,vol. 7, pp 44-55, 2013.

[7] P Arrigo, F. Giuliano, F. Scalia, A Rapallo and G.Diamiani, “Identification of a new motif on nucleic acid sequence data using Kohonen’s self-organizing map”, Computer Applied Bioscience, vol. 7, pp 353-357, 1991.

[8] T. Kohonen, “Self organizing formation of topologically correct feature maps”, Biol. Cybern, vol. 43, pp 59-69, 1982.

[9] Gabriela Czibula and Istvan Gergely Czibula,”Unsupervised restructuring of object oriented software Systems using Self organizing feature maps”, International Journal of Innovative Computing, Information and Control,vol.8 , pp 1689-1704, 2012.

[10] J. Vesanto, J. Himberg, E. Alhoniemi and J. Perhankangas, “ Self Organizing Map in Matlab: The Som ToolBox”, DSP conference, vol 2010, pp 35-40, 1999.

[11] Jesna M. and Kumudh Riamond, ”MR Brain Image Segmentation based on Principle Component Analysis and Self Organizing Map”, International Journal for Research in Applied Science and Engineering Technology, vol.2, pp 161-167, 2014.

[12] D Jander Moreira and Luciano Da Fontoura Costa, “Neural-based color image segmentation and classification using self-organizing maps”, Proceedings of the conference on graphics patterns and images, pp 47-54,1996.