# MEMBERSHIP FUNCTION – ARTMAP NEURAL NETWORKS

PETER SINČÁK*[1], MARCEL HRIC[2] AND JÁN VAŠČÁK**[2]

[1] *Siemens AG, Vienna PSE Department,*
*ECANSE working group,*
*Gudrunstrsse II, 1100 Vienna, Austria*
*Peter.Sincak@tuke.sk*

[2] *Center for Intelligent Technologies,*
*Faculty of EE and Informatics Technical University,*
*Letná 9, 04001 Košice, Slovak Republic*
*hric@neuron-ai.tuke.sk, Vascak@tuke.sk*

(Received 28 October 2001; revised manuscript received 1 August 2002)

**Abstract:** The project deals with the application of computational intelligence (CI) tools for multispectral image classification. Pattern recognition scheme is a global approach where the classification part is playing an important role to achieve the highest classification accuracy. Multispectral images are data mainly used in remote sensing and this kind of classification is very difficult to assess the accuracy of classification results. There is a feedback problem in adjusting the parts of pattern recognition scheme. Precise classification accuracy assessment is almost impossible to obtain, being an extremely laborious procedure. The paper presents simple neural networks for multispectral image classification, ARTMAP-like neural networks as more sophisticated tools for classification, and a modular approach to achieve the highest classification accuracy of multispectral images. There is a strong link to advances in computer technology, which gives much better conditions for modelling more sophisticated classifiers for multispectral images.

**Keywords:** pattern recognition principles, classifier design, classification accuracy assessment, contingency tables, backpropagation neural networks, fuzzy BP neural networks, ART and ARTMAP neural networks, modular neural networks

## 1. Introduction

Image processing is an important part of modern technology and its applications include environmental monitoring, very important for the future of humankind. Image processing can be the most important part of this system and determines the quality

* P. Sinčák is on temporary leave from Center for Intelligent Technologies at TU Košice Slovakia and his stay as guest professor at Siemens, Vienna is supported by Maria-Curie fellowship from EU 5FP activity

** On leave at Hirota Laboratory, Tokyo Institute of Technology Japan, vascak@hrt.dis.titech.ac.jp

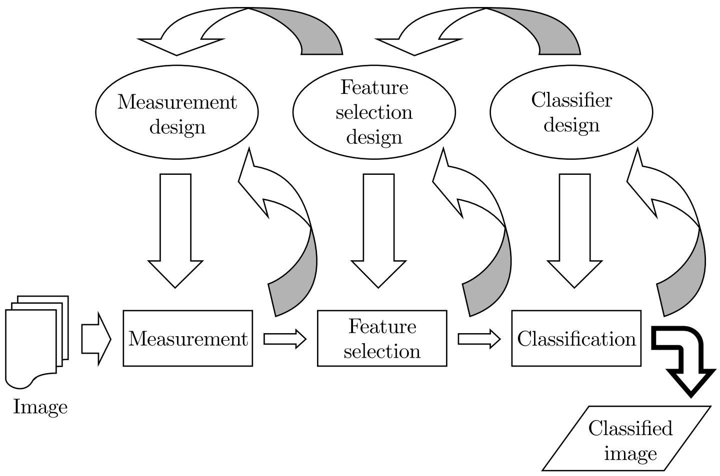


**Figure 1.** Basic principles of pattern recognition

accuracy of information retrieval from the images. Generally, we have to take into consideration the basic principles of pattern recognition, where classification plays an important role. These principles are presented in Figure 1. As it is clear from the above figure, the determination of feature space plays a key role in overall pattern recognition principles. Feature selection and classification are based on this information and the level of ability to approximate the nonlinear discrimination function determines the success of a classification procedure. A very important aspect of this procedure is the accuracy assessment of classification results. Contingency analysis is a proper way of accuracy assessment and it reflects the multiclass results of the classification procedure [1–4].

Important aspects of the pattern recognition principle are the adaptability of the overall procedure and the equal importance of all steps in the procedure, including feature space determination and a feature selection procedure. This pattern recognition principle reveals the following questions:

- Is it more useful to "invest" more time in the best feature selection procedure?
- Or is it more important to „invest" more time in development of the most sophisticated classification procedure for any type of features?

These two questions play a key role in the future of research in the domain of classification. Certainly, these two problems are closely connected and usually feature selection techniques assume a specific type of classifier. We can find a number of techniques dealing with the first approach and a number of approaches dealing with the second problem. Basically, the first approach is based on the assumption that a simple and fast classification procedure is used while the second problem is related to the assumption of simple and almost raw data or data with minimal preprocessing procedures. Certainly, a combination of both of these approaches could be fruitful and yield good results for real-world application.

## 2. Motivation of the project

Classification is mapping from a feature space into a space of classes. It is often very difficult to determine whether a certain point in a feature space belongs to a certain class. Therefore, an approach based on fuzzy sets has many advantages in reducing misclassification in the results. Sometimes it is more convenient to have results in the form of transparent information concerning relations of the observed point in the feature space to all classes of interest. Instead of a crisp classifier output, we can be more satisfied with outputs based on fuzzy sets, *viz.* on values of membership functions of the observed input to fuzzy clusters and fuzzy classes. The notions of fuzzy clusters and fuzzy classes are described below. The motivation is to provide the end-user with a smaller number of misclassifications and higher readability of the classification results. The output of these classification results is a vector of values describing relation of the input to each class of interest. The desire is to have a highly parallelized tool with an incremental learning ability similar to that of an ARTMAP family neural network [5–8].

## 3. Description of the method

The project is based on the assumption that data in feature space are organized in fuzzy clusters. Fuzzy cluster $A$ is considered as a fuzzy set $A$ in multidimensional feature space representing a set of ordered couples *e.g.*:

$$A \in \{[x_1,\mu_A(x_1)],\ldots,[x_n \mu_A(x_n)]\}, \tag{1}$$

where $A$ is a fuzzy set and $\{\mathbf{x},\mu_A(\mathbf{x})\}$ are ordered couples, $\mathbf{x}=[x_1,\ldots,x_n]^T$ being a point in multidimensional feature space and $\mu_A(\mathbf{x})$ – a value of the membership function of $\mathbf{x}$ to the fuzzy cluster (set) $A$. There are many fuzzy clusters in the feature space and a certain set of fuzzy clusters create a fuzzy class. Fuzzy class is the union of fuzzy clusters belonging to a considered class defined by a training set *e.g.*:

$$CL = \left\{ \bigcup_{i=1}^{n} A_i \right\}. \tag{2}$$

Generally, we can consider a fuzzy class as a set of fuzzy sets $A_i$ representing the variety of the numerical representation of the class. Relation between $\mu_{CL}(\mathbf{x})$ and $\mu_A(\mathbf{x})$ must be as follows:

$$\mu_{CL}(\mathbf{x}) = \max_{i=1,2,\ldots,n} (\mu_{A_i}(\mathbf{x})), \tag{3}$$

where $A_i$ is a fuzzy cluster which belongs to class $CL$ and $n$ is a number of fuzzy clusters creating class $CL$. The membership function is considered as:

$$\mu_A(\mathbf{x}) = \frac{1}{\left[1+e^{\frac{\mathbf{x}-mc}{\sigma}}\right]^F}, \tag{4}$$

where $F=1$, $mc$ is the mean value and $\sigma$ is variance of a Gaussian type of membership function. The MF-ARTMAP is intended as the tool to calculate values of membership functions of $\mathbf{x}$ to each class of interest in the feature space.

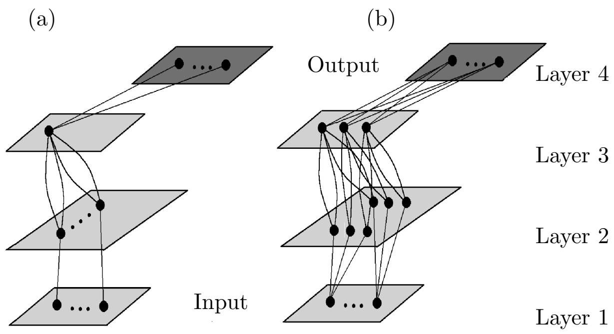


**Figure 2.** Basic topology of MF-ARTMAP neural networks: (a) MF-ARTMAP with one cluster, (b) MF-ARTMAP with three clusters

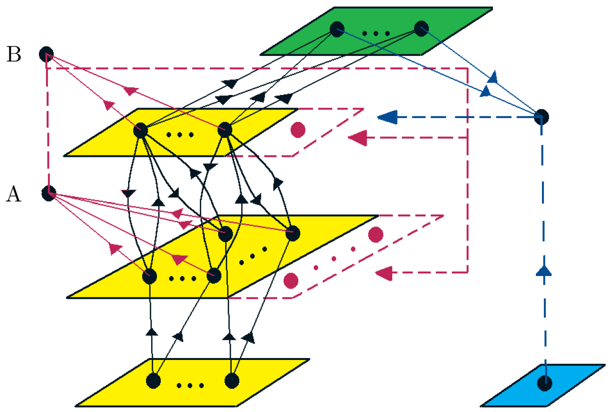


**Figure 3.** General MF-ARTMAP topology with a dynamic number of neurons in the 2nd and 3rd neural network layer

### *3.1. Description of the neural network topology*

Topology of MF-ARTMAP is based on an architecture similar to that of ARTMAP. In Figure 2, we can see two MF-ARTMAP neural networks with 4 neural layers. The input layer is mapping the input into the comparison layer, where the comparison is made between the input pattern and mean values of the existing clusters. In Figure 2a, an initial situation is shown, where only one cluster is identified, while in Figure 2b, 3 clusters are already revealed, so that it is tested whether the input pattern belongs to any of the clusters. In this case, the second layer is changing dynamically according to the number of clusters in the 3rd layer.

Thus, the 2nd and 3rd layers are extending according to the number of clusters found in the feature space. The recurrent connection between the 2nd and 3rd layer is the encoding of mean value and mean square deviation associated with a particular membership function based on a Gaussian type of representation. The 4th layer represents the mapfield as a part of a neural network, whose role is to integrate the clusters into a resulting class. The input to the mapfield is from the 3rd layer as well as from outside the neural network, as associated output into the overall MF-ARTMAP neural networks. A more detailed topology is shown in Figure 3, where a situation with more clusters is represented. Basically, the neural layers in the MF-ARTMAP can be listed as follows:

Layer 1 – an input mapping layer, with the number of neurons equal to "$n$", where $n$ is the dimensionality of the feature space,

Layer 2 – a comparison layer, with the number of neurons equal to "$n \times nc$", where $nc$ is the number of clusters identified in the recognition layer,

Layer 3 – the recognition layer, with the number of neurons equal to "$nc$", where $nc$ is the number of clusters,

Layer 4 – the mapfield layer, with the number of neurons equal to "$M$", where $M$ is the number of classes for the classification procedure.

A more detailed description of the MF-ARTMAP is presented in Figure 3. An important function is performed by neurons A and B, which reset and freeze some of the procedures during training.

These neurons are fed from the comparison layer with values which represent a difference between input patterns and mean values of the input feature values. If the maximum of these differences is greater than the threshold, then a new cluster is set up immediately. If not, there is a chance that the input pattern belongs to one of the clusters identified in the past procedures. Then, training and updating of the recurrent synapses is under way to adapt the shape of the membership function. The adaptation procedure is described in the following section. More information can be found in [9].

### 3.2. Parallel MF-ARTMAP

The notion of a modular neural network has been known for many years. It is a very promising idea of solving complex problems by their distribution into more sub-problems which are easier to solve. Similar approach can be found in [10].

Basically, the "divide and conquer" principle is usually used in modular neural networks. There are some difficult questions about separability problems and discrimination of hyper-planes determination. The key problem is as follows: "Is it easier to separate one particular class from the feature space or to identify more classes?"

In fact, the question is about the difficulties of dichotomous classification as compared with the multi-class approach. The first impression may be that dichotomous classification is always easier than the multi-class approach, but it is very difficult to conclude in general. To investigate these ideas, a parallel ARTMAP approach was designed and tested. The basic philosophy of this approach is illustrated in Figure 4. The MF-ARTMAP is suitable for solving the conflict in this approach because the values of membership functions to fuzzy clusters are good indicators of conflict solution among more experts, as it is indicated in Figure 4. The basic advantages of Parallel MF-ARTMAP are as follows:

1. Ultra-fast learning abilities on highly parallel systems, *e.g.* PC-farms. This feature can be very useful for large databases, with easier handling of larger amount of data.

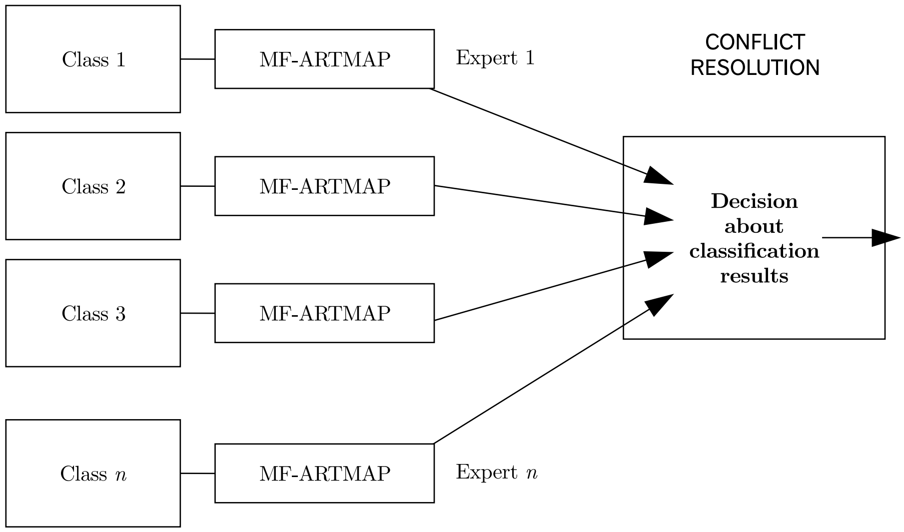


**Figure 4.** The basic philosophy of Parallel ARTMAP

2. Easy and comfortable extension of classes of interest by adding a new expert network and training on the new class training data. This is very interesting in case of frequent additions of new classes to the list of classes of interest.
3. Easy identification of class "unknown" by measuring the membership function value of the unknown input to the fuzzy classes. If the value is lower than the given value – then it is rejected and proclaimed class "unknown". This feature is very important when large amounts of data with many classes are considered.
4. Easy readability of fuzzy classes as unions of fuzzy clusters and identification of their basic parameters.

On the other hand, the main disadvantage of Parallel MF-ARTMAP is *the necessity of more parameter determination considering starting values of vigilance parameters for each expert network separately.* This disadvantage can be avoided by designing a meta-controller setting up all the parameters for the overall complex of experts, *e.g.* a fuzzy system.

## 4. Experimental results

Experiments using benchmark and real-world data have been made as a part of this project. The aim was a comparative analysis of the previously known CI systems with those modified or developed during the research, namely MF-ARTMAP, Parallel ARTMAP. Basically, the same real-world data were used and so the comparison could be made assuming the same training and testing data, which means that the same amount of knowledge was used for the classification procedures.

### 4.1. Accuracy assessment

Accuracy assessment was made using contingency table analysis. A contingency table was used in a basic comparison study between all the methods which were investigated and developed. Some details about contingency table analysis for accuracy assessment of classification results can be found in [11].

**Table 1.** Results on the "circle in a square" dichotomous classification

| Predicted class | Actual class | | | |
|---|---|---|---|---|
| | MF-ARTMAP | | Parallel MF-ARTMAP | |
| | A | B | A | B |
| A' | 97.62 | 2.92 | 98.14 | 1.46 |
| B' | 2.38 | 97.08 | 1.86 | 98.54 |

**Table 2.** Results on the "double spiral" dichotomous classification

| Predicted class | Actual class | | | |
|---|---|---|---|---|
| | MF-ARTMAP | | Parallel ARTMAP | |
| | A | B | A | B |
| A' | 87.54 | 10.76 | 87.96 | 7.86 |
| B' | 12.46 | 89.24 | 12.04 | 92.14 |

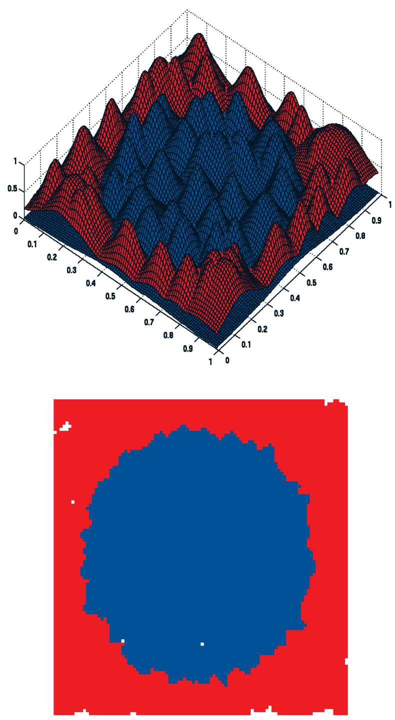

**Figure 5.** Classification of circle in the square (Parallel MF-ARTMAP, MF-ARTMAP) according to Table 1

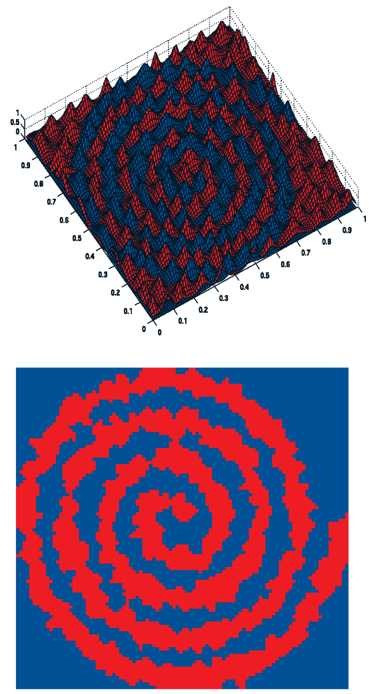

**Figure 6.** Classification of double spiral (Parallel MF-ARTMAP, MF-ARTMAP) according to Table 2

## *4.2. Experiments on benchmark data*

There were 2 benchmark data used for testing classification results for comparative purposes. The circle in a rectangle (Figure 5) and the double spiral problems (Figure 6) were used for dichotomous classification testing. These two benchmark

**Figure 7.** Original image – highlighted areas were classified by expert (A – urban area, B – barren fields, C – bushes, D – agricultural fields, E – meadows, F – forests, G – water)

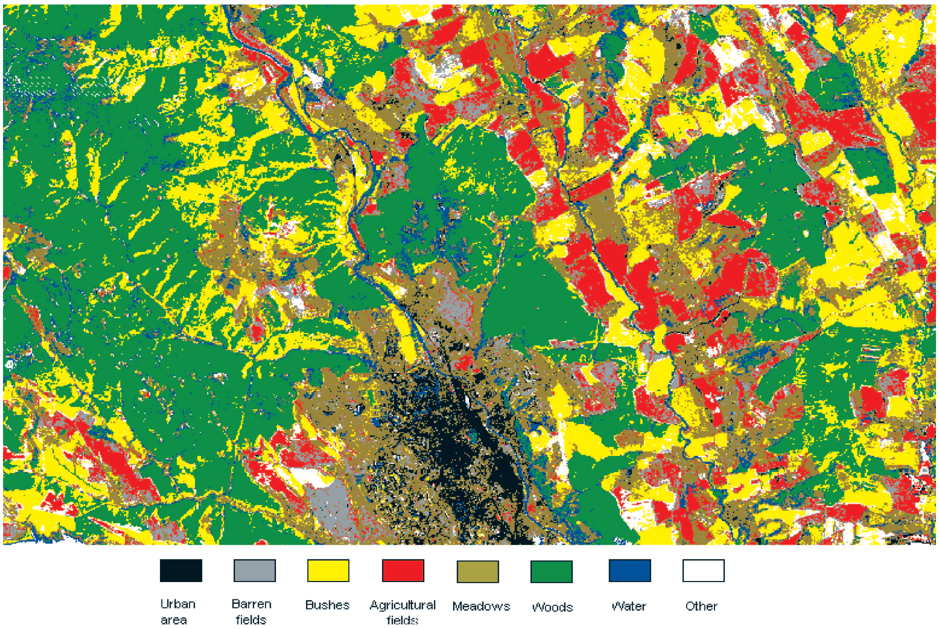


**Figure 8.** Classification results on Landsat TM data using Parallel MF-ARTMAP approach

data are used mostly for estimating the level of sophistication of classifiers. If results on these benchmark data are good, there is a good assumption that the classifier will be successful in the main applications. Tables 1 and 2 present the classification results obtained by the use of testing data from the selected benchmarks.

TQ107C-I/50 10 X 2003 BOP s.c., http://www.bop.com.pl

**Table 3.** Contingency table of Landsat TM classification on the test sites for MF-ARTMAP

| Predicted | Actual | | | | | | |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G |
| A' | 95.51 | 0.21 | 0.00 | 0.00 | 2.76 | 0.00 | 4.23 |
| B' | 0.00 | 83.16 | 0.00 | 3.01 | 7.18 | 0.00 | 0.00 |
| C' | 0.00 | 0.00 | 100 | 0.00 | 1.66 | 0.00 | 0.00 |
| D' | 0.00 | 11.99 | 0.00 | 96.66 | 1.10 | 0.00 | 0.00 |
| E' | 3.85 | 4.72 | 0.00 | 0.22 | 87.29 | 0.17 | 7.04 |
| F' | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 99.49 | 5.36 |
| G' | 0.64 | 0.00 | 0.00 | 0.11 | 0.00 | 0.34 | 83.10 |

**Table 4.** Contingency table of Landsat TM classification on the test sites for Parallel MF-ARTMAP

| Predicted | Actual | | | | | | |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G |
| A' | 94.87 | 0.83 | 0.00 | 0.00 | 4.68 | 0.00 | 5.41 |
| B' | 0.00 | 84.30 | 0.00 | 3.77 | 6.43 | 0.00 | 0.00 |
| C' | 0.00 | 0.00 | 99.03 | 0.00 | 0.00 | 0.00 | 0.00 |
| D' | 0.64 | 9.71 | 0.00 | 96.01 | 1.17 | 0.00 | 1.35 |
| E' | 4.49 | 4.96 | 0.97 | 0.22 | 87.72 | 0.17 | 4.05 |
| F' | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 99.74 | 0.00 |
| G' | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.09 | 89.19 |

### *4.3. Experiments on real-world data*

Experiments were conducted on benchmark and real-world data. Basically, the behaviors of the methods were observed on multispectral image data with the aim of obtaining the best classification accuracy on the test data subset. The Košice data consist of a training set of 3164 points in the feature space and of a test set of 3167 points of the feature space. A point in the feature space has 7 real-valued coordinates of the feature space normalized into the (0,1) interval and 7 binary output values. The class of a fact is determined by the output which has a value of one; the other six output values are zero. The data represent 7 attributes of the color spectrum sensed from the Landsat satellite. The representation set was determined by a geographer and was supported by a ground verification procedure. The main goal was land-use identification with the most precise classification procedure to achieve accurate results. The image was taken in the eastern Slovakia region, particularly in the vicinity of the city of Košice. There were seven classes of interest picked up for the classification procedure, as shown in Figure 7. The results are presented in Tables 3 and 4. In Figure 8 classified original image by Parallel MF-ARTMAP approach is shown.

## 5. Conclusion

The paper presents further research on the Parallel MF-ARTMAP approach to the classification procedure. The advantage of this approach is higher readability of the neural network in providing the values of membership functions of input to fuzzy clusters or fuzzy classes identified by this approach. Results of these neural networks are comparable with MF-ARTMAP [9] on benchmark and real-world data and, additionally, provide more useful information about the measure of membership in all the fuzzy classes and fuzzy clusters discovered in the feature space. This seems to be an interesting advantage of this approach. Further work is under way to apply Parallel MF-ARTMAP to financial data, for classification purposes.

## References

[1] Duda R and Hart P 1973 *Pattern Classification and Scene Analysis*, Wiley, New York

[2] Werbos P 1974 *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, PhD Thesis, Harvard University

[3] Ocelikova E and Klimesova D 2000 *Proc. 4th Int. Scientific Technical Conference Process Control – ŘÍP 2000*, Pardubice, Czech Republic, p. 108

[4] Klimesova D and Ocelikova E 2001 *Proc. Int. Conf. Agrarian Perspectives X-Sources of Sutainable Economic Growth in the Third Millenium. Globalisation versus Regionalism*, Prague, Czech Republic, pp. 67–74

[5] Carpenter G A, Gjaja M N, Gopal S and Woodcock C E 1997 *IEEE Trans. Geoscience and Remote Sensing* **35** (2) 308

[6] Grossberg S 1976 *Biological Cybernetics* **23** 187

[7] Carpenter G A, Milenova B L and Noeske B W 1998 *Neural Networks* **11** (5) 793

[8] Williamson J R 1996 *Neural Networks* **9** 881

[9] Sinčák P, Kopčo N, Hric M and Veregin H 2000 *Proc. 6th Int. Conf. on Computational Intelligence, IIZUKA*, Iizuka, Japan (1)

[10] Rutkowska D and Nowicki R 2000 *Int. J. Appl. Math. and Comput. Sci.* **10** (4) 675

[11] Sinčák P, Veregin H and Kopčo N 2000 *Geocarto Int.* **15** 11