An Approach for Invariant Clustering and Recognition in Dynamic Environment

Andrey Gavrilov, Sungyoung Lee

Computer Engineering Department, Kyung Hee University,
1 Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do, 446-701,
Republic of Korea

Abstract - **An approach for invariant clustering and recognition of objects (situation) in dynamic environment is proposed. This approach is based on the combination of clustering by using unsupervised neural network (in particular ART-2) and preprocessing of sensor information by using forward multi-layer perceptron (MLP) with error back propagation (EBP) which supervised by clustering neural network. Using MLP with EBP allows to recognize a pattern with relatively small transformations (shift, rotation, scaling) as a known previous cluster and to reduce producing large number of clusters in dynamic environment, e.g. during movement of robot or recognition of novelty in security system.**

I. INTRODUCTION

Most important tasks for actions of mobile robot in unknown environment using behavior based approach [1] are the clustering and recognition of objects and situations. Application of Adaptive Resonance Theory (ART) [2] (in particular ART-2) is rather appropriate for solving these tasks, because this model combines properties of plasticity and stability, and also it does not require a priori knowledge of the fixed quantity of necessary classes. Many different modifications of this model and its combinations with other neural networks are known [3], [4, [5] and others. However, most of them (except FANNC [5]) are supervised learning and demand a teacher. Thus, the most important feature of model ART is refused. As to model FANNC, this one aims to classify static images and is not oriented on recognition of novelty in flow of images many of which may be similar and differ only by some gradual transformations. So we focus on model ART-2 to provide fast unsupervised learning.

However, model ART-2 assumes usage of only one layer of neurons (not including entry, associated with sensors). It results in that the neural network works only with the metrics of primary features, e.g. for visual images these features are pixels. At that similarity between images (for classification or creation of a new cluster - output neuron), are calculated usually using Euclidean distance. As result, even any small transformation of input vector can cause recognition of it as belonging to new cluster. It leads to the result that the model ART-2 is nearly never used for real applications. For example, considering the clustering and pattern recognition of a mobile robot, [6], [7] it is required to recognize the object in different foreshortenings which may allocated in different parts of a field of vision, i.e. recognition should be invariant concerning transformation of the input pattern, such as shifts, rotations and others.

The ability of invariant recognition of objects (situations) in dynamic environment is one of most important capabilities of natural brain [8]. This ability is closely connected with another ability which is to detect novelty. These are "two sides of one same medal".

In this paper we suggest and investigate one hybrid model of neural network based on ART-2 and multi layer perceptron (MLP) with error back propagation (EBP) training algorithm [9]. In this model we keep the unsupervised learning and remove one major disadvantage of ART – the sensitivity to transformations of input patterns. Multi layer perceptron provides preprocessing of patterns for invariant recognition because its hidden layers form secondary features during learning. It could be said that in MLP each hidden layer provides conversion of any feature space to another one.

There are many other approaches to achieve invariant recognition by neural networks, for example proposed in [10], [11], [12], [13], [14], [15], [16]. But each of them is either too complex or specialized for determined any kind of images and transformations.

We suggest potentially universal approach for invariant recognition which can be implemented in real time systems, because it not requires long time processing as in usual applications of EBP.

For the first time the paradigm of MLP-ART2 model was proposed by us in [17], [18]. In that paper the idea was suggested and some previous experiments and problems were shown. In current paper this paradigm and algorithm of model get continue developments. Some new ideas, experiments and conclusions are discussed.

II. MAIN CONCEPTS AND ALGORITHM OF HYBRID NEURAL NETWORK MLP-ART2

In our suggested model the first several layers of neurons are organized as MLP. Its outputs are the inputs of model ART-2. MLP provides conversion of primary feature space to secondary feature space with lower dimension. Neural network ART-2 classifies images and uses secondary features to do it. Training of MLP by EBP (with limited small number of iterations) provides any movement of an output vector of MLP to centre of recognized cluster of ART-2 in feature space. In this case the weight vector (center) of recognized

cluster is desired output vector of MLP. It could be said that the recognized class is a context in which system try to recognize other images as previous, and in some limits the system "is ready to recognize" its by this manner. By other words neural network "try to keep recognized pattern inside corresponding cluster which is recognizing now".
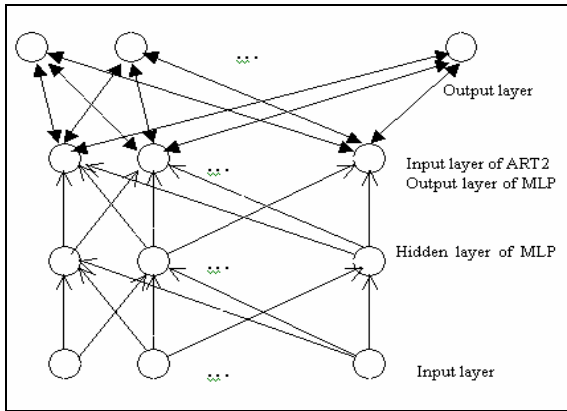

Fig. 1. Structure of hybrid neural network.

Action of the suggested model is described by the following unsupervised learning algorithm:

1. In MLP let the weights of connections equal to $1/n$, where n is quantity of neurons in the previous layer (number of features for first hidden layer). The quantity of output neurons N_{out} of ART-2 is considered equal zero.

2. The next example from training set is presented to inputs of MLP. Outputs of MLP are calculating.

3. If $N_{out}=0$, then the output neuron is formed with the weights of connections equal to values of inputs of model ART-2 (the outputs of MLP).

4. If $N_{out}> 0$, in ART-2 the algorithm of calculation of distances between its input vector and centers of existing clusters (the weight vectors of output neurons) is executing using Euclidian distance:

$$d_j = \sqrt{\sum_i (y_i - w_{ij})^2}$$

Where: y_i – i^{th} feature of input vector of ART-2, w_{ij} – i^{th} feature of weight vector of jth output neuron (the center of cluster). After that the algorithm selects the output neuron-winner with minimal distance. If the distance for the neuron-winner is more than defined a vigilance threshold or radius of cluster R, the new cluster is created as in step 3.

5. If the distance for the neuron-winner is less than R, then in model ART-2 weights of connections for the neuron-winner are recalculated by:

$$w_{im} = w_{im} + (y_i - w_{im})/(1 + N_m)$$

Where: N_m – a number of recognized input vectors of mth cluster before. In addition, for MLP a recalculation of

weights by algorithm EBP is executing. In this case a new weight vector of output neuron-winner in model ART-2 is employed as desirable output vector for EBP, and the quantity of iterations may be small enough (e.g., there may be only one iteration).

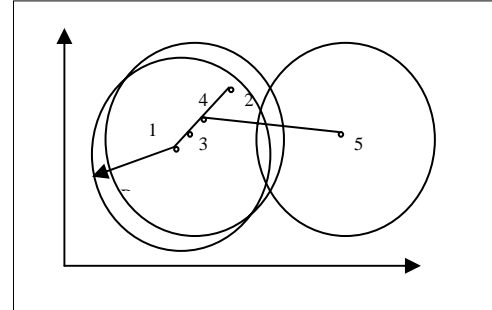6. The algorithm repeats from step 2 while there are learning examples in training set.

Fig. 2. Explanation of action of hybrid model.

Action of the suggested model is explained in a figure 2. Here we see the space of secondary features in which by points are represented output vector of perceptron (input vector of model ART-2) and centers of clusters. In this figure the following points are represented:

1) - a new (first) image for which new cluster with radius R is created,

2) - the next image recognized as concerning to this cluster,

3) - the new centre of a cluster calculated in step 5 of algorithm,

4) - a new output vector of perceptron, attracted to centre of a cluster as a result of executing of algorithm "error back propagation",

5) - the new image recognized as inhered to other cluster.

Note that in this algorithm EBP aims at absolutely another goal different from that in usual MLP-based systems. In those systems EBP reduces error-function to very small value. But in our algorithm EBP is needed only for some decreasing distance between actual and desirable output vectors of MLP. So in our case the long time learning of MLP is not required.

Note that EBP and forming of secondary features are executed only when image "is captured" by known cluster. So selection of value for vigilance threshold is very important. Intuitively obvious that one must be depending on transformation speed of input images and may be changed during action of system.

In our experiments we used three variants for calculation of vigilance threshold:

1) it is fixed value selected empirically,

2) it is calculated for every image by formulas $S/(N_a^2)$, where S – average input signal, N_a – number of output neurons of MLP (input neurons of ART-2),

3) it is calculated as kD_{min}, where D_{min} – minimal distance between input vector of ART2 and weight vector in previous image, $k>1$ – coefficient.

For experiments we developed program for simulation in which we process series of visual images. Every image consists of 100x100 pixels. The model MLP-ART2 consists of one hidden layer of MLP. Program is implemented in Delphi.

III. EXPERIMENTS

A.. *Experiments with Images of Moving Objects*

We use series of multi-colored images for experiments which obtained from video. This series simulates the sight of mobile robot during its movement. There are images of shifting couple of chairs. Figure 3 shows the images 1, 8 and 15 (last) of this series.

Fig. 3. First image of series from video

Figures 4 and 5 show results of clustering and recognition at repeating of this series four times. In these experiments we used following parameters of neural network:
- the number of hidden neurons of MLP – 20,
- the number of output neurons of MLP (or input neurons of ART-2 – 10,
- exponential sigmoid with parameter $a=1$,
- the number of iterations of EBP - 1,
- fixed value of vigilance threshold $R=0,05$.

Figure 4 is the evidence of stability of our model at constant value of vigilance threshold. Cluster 2 is corresponding to situation when the robot sees only one chair; and cluster 3 is corresponding to situation when robot sees basically already empty wall. Figure 5 shows distance between input vector of ART-2 and center of recognized cluster 1 for first 7 images after 4 times of repeated series.

In case of variable vigilance threshold the stability is less. Figure 6 shows results with calculation of one by method 3 with $k=4$ and number of iterations of EBP equals to 5.
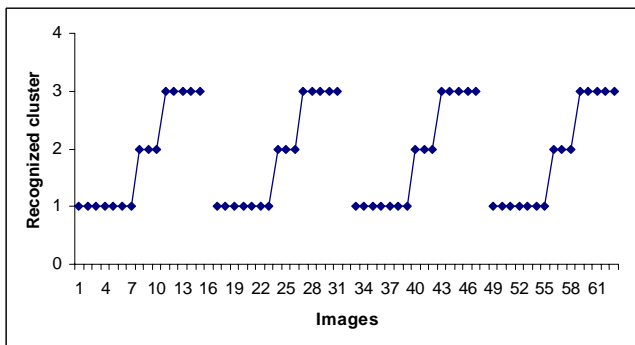

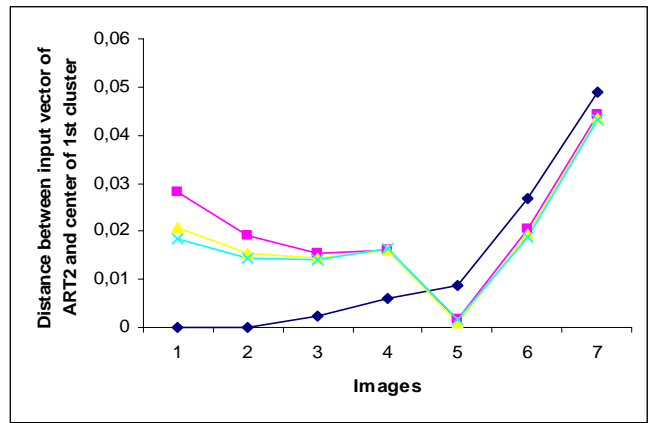Fig. 4. Recognized clusters at repeating of series of images

Fig. 5. Distance between recognized cluster 1 and input vector of ART-2 (output vector of MLP).
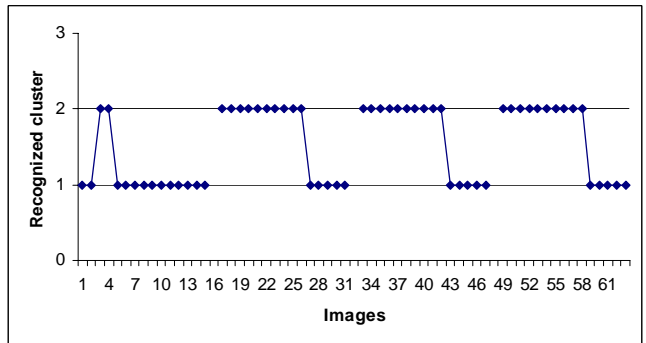

Fig. 6. Recognized clusters at usage of calculation of vigilance threshold by method 3 with k=4

Note that in repetitions 2, 3 and 4 the results of recognition are more intuitively well-taken than in first time.

Figure 7 shows results of clustering and recognition in case when vigilance threshold is calculated by method 3 with $k=2$ and number of iterations of EBP equals to 1.
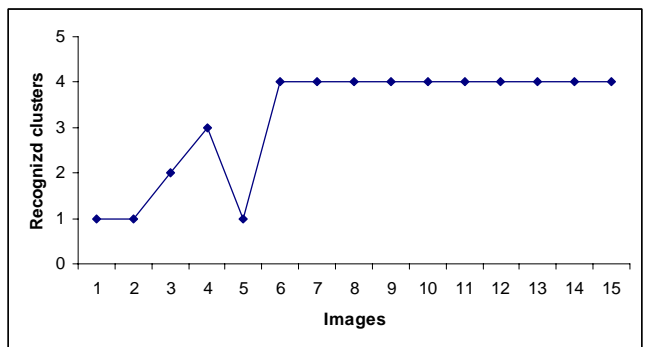

Fig. 7. Recognized clusters at usage of calculation of vigilance threshold by method 3 with k=2.

Figure 8 shows distance between input vector of ART-2 and center of cluster 4 during its recognition for same parameters and number of iterations of EBP equals to 1, 3, 5 and 10. In this figure lower lines are corresponding to large number of iterations of EBP. Thus we can see that the

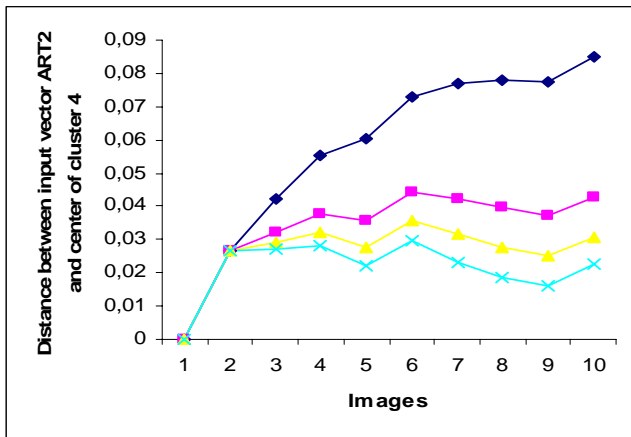distance is decreasing in accordance with increasing of number of iterations.

Fig. 8. Distance between input vector of ART-2 and center of recognized cluster 4 with different number of iterations of EBP

B. Experiments with Series of Faces

For these experiments we use collection of faces from Olivetti and Yale. Series of images used in experiments are shown in figure 9. This series consists of set of gray-scaled pictures of certain person with different expressions (from 1 to 8), with different illumination (from 9 to 11), with transformation by turning (from 12 to 18), another 3 faces (from 19 to 21), faces of last person with different expressions and position (from 22 to 29), repetition of face 19 (30).

In these experiments following parameters of model are used:

- the vigilance threshold is calculated for every image by method 2, described above,
- the number of iterations of EBP - 1,
- the number of inputs of ART-2 = 10,
- the number of hidden neurons of MLP – 20,
- exponential sigmoid with parameter a=1,

Figure 10 shows results of clustering and recognition of two repeating of series. We can see from this picture that the model demonstrates stable results. Figure 11 partially explains the result shown in Figure 10. The images 9, 10, 11 with other illuminations in contrast to previous images are difficult for recognition and cause creation of new clusters. Images 19-21 are recognized as new persons and images 22-29 - as 2 different faces in accordance with its mimic and position features. Image 30 is recognized like previously looked face 19.

Figure 12 shows distance between input vector of ART-2 and center of recognized cluster 1 with different number of iterations of EBP: 1, 3, 5 and 10. The line beginning from lowest point is corresponding to value 1, and the line from highest point is corresponding to value 3 and further increasing number of iterations causes decreasing of first point and increasing of last point in this graphic.
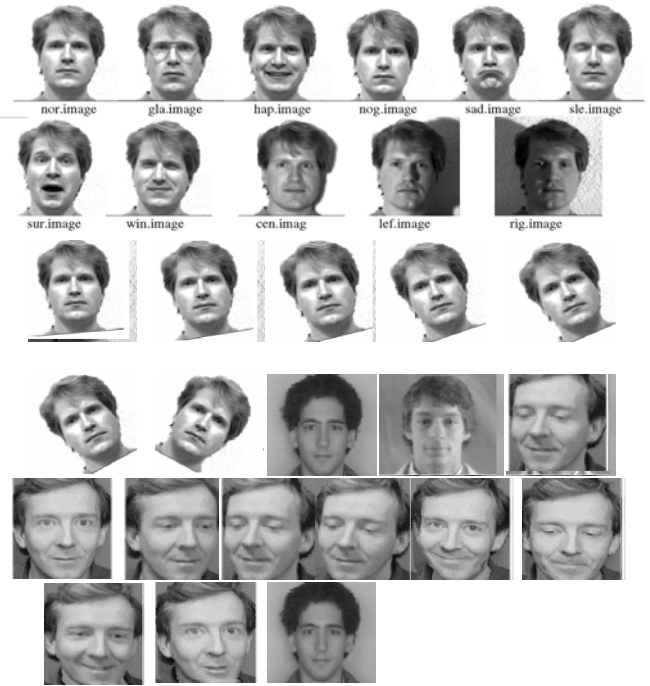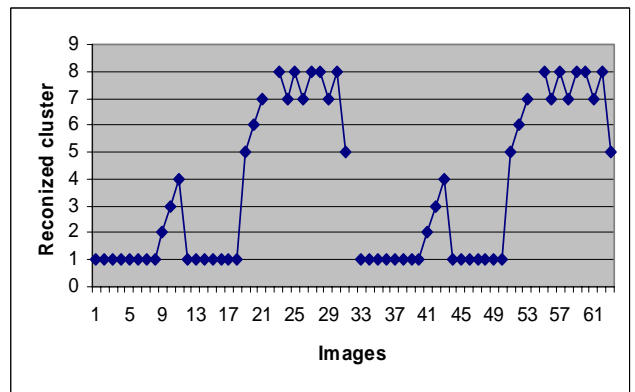

Fig. 9. Series of faces

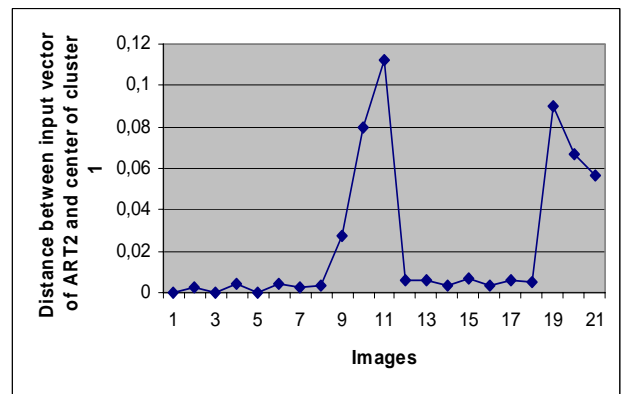Fig. 10. Recognized clusters for 2 repetitions of series of faces

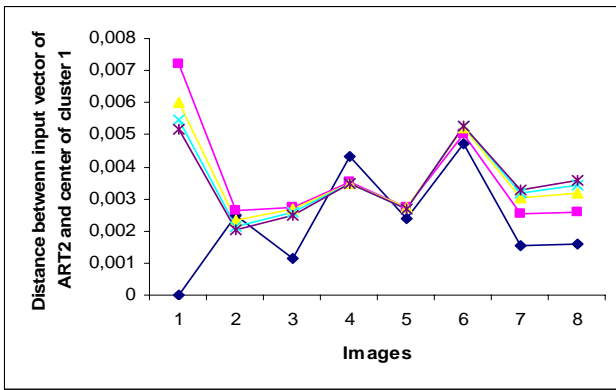Fig. 11 Distance between input vector of ART-2 and center of cluster 1

Fig. 12. Distance between input vector of ART-2 and center of recognized cluster 1 for first 8 images of series for different number of iterations for EBP

Figure 13 shows results of clustering and recognition in case with calculation of vigilance threshold by methods 3 with $k=4$, number of iterations equals to 5 and number of repetitions of series equals to 4.

From these results we can see that in this case every new repetition of series improves accuracy of recognition. At beginning all images of first person with other illumination are recognized as one cluster 2. Almost all images of other persons are recognized as one cluster 3 (may be interpreted as "other person"). In next repetition different variants of illuminations are recognized as different clusters. In third repetition other persons are recognized as different persons. But in this case one of these persons is recognized as person 1 with other illumination (cluster 6). In last repetition this single mistake is removed. It could be said that repetition causes more detailed clustering and recognition that is very similar to well known features of human learning.

IV. DISCUSSIONS

The demonstrated above and other obtained results of experiments show that proposed and investigated model MLP-ART2 of hybrid neural network demonstrates adaptation to relatively small transformations of visual image and capability to recognize new images in flow of ones. This adaptation coexists with keeping ability to recognize previously occurred images. For all that this model demonstrates some features of perception of humans, such as "we see what we expect to see" [8] and sometimes effect of some difficulty to recognize following new images immediately after adaptation to recognition of any previous one (new different images are recognized as one cluster). Also experiments show stability and capability to improve accuracy (in human-like respect) of recognition and clustering during repetition of same series of images.

The suggested hybrid model of the neural network can be used in the mobile robot when it is necessary to watch sequence of images visible by the robot during its movement, and to extract its novelty, i.e. essential changes in the scene visible by the robot. Also this model may be used in security systems for recognition of new objects or persons in sight of

camera. Another possible application of this approach is recognition of novelty and dangerous situations in different monitoring systems, e.g. in health care systems.
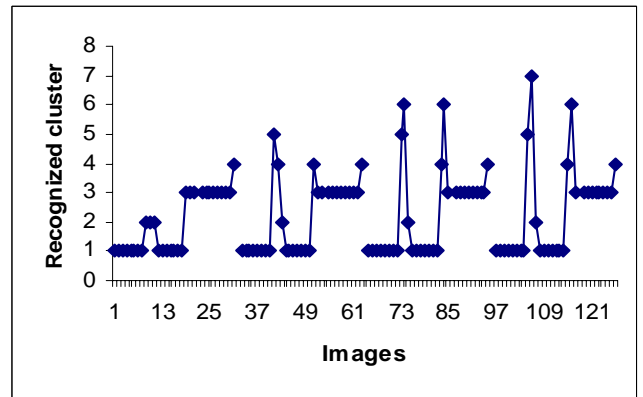

Fig. 13. Recognized clusters at calculation of vigilance threshold by method 3 with $k=4$

Modification of this algorithm may be algorithm in which the quantity of created clusters is limited. In this case, if the quantity of clusters (output neurons) has achieved a limitation, it is needed to decide what to do with images which are not recognized, i.e. which cannot be related to any existing cluster. In this case following method may be offered. At first, increasing of parameter R (radius of clusters) and then trying to apply algorithm of recognition again until the new image will not be related to one of clusters. After that, it is necessary to reduce quantity of clusters (output neurons) to unify clusters with the centers which are occurred in one cluster with new value R, and to change weights of connections between outputs of MLP and outputs neurons-clusters.

Also one modification of this algorithm may be algorithm of training with the teacher in which before creating new cluster the system requests teacher what to do in order to create a new cluster or to increase vigilance threshold R. Moreover in this case may be not only to ask from user but also to employ any additional procedure for recognition of novelty of the image.

Similar approach for building hybrid neural network from supervised and unsupervised neural networks may be applied to self-organized maps [19].

V. CONCLUSIONS

In this paper we suggested and investigated hybrid model of neural network MLP-ART2 for unsupervised clustering and invariant recognition of images in dynamical unknown environment. The suggested hybrid model of the neural network can be used in the mobile robot and security systems when it is necessary to watch sequence of the visible images and to recognize similar or novel objects (situations).

This approach may be applied to self-organized maps of T.Kohonen.

The following further researches of the suggested hybrid model of the neural network are planned:

- continue research of influence of MLP and ART-2 parameters on action of the neural network,

- testing and investigation of the suggested model on simulated mobile robot and the real robot,

- development of hierarchical system based on proposed hybrid neural network which may be basis for building of model of mind,

- development of hybrid architecture of intelligent system based on proposed model and knowledge based approach for perception and processing of multi-modal information.

REFERENCES

[1] R.A. Brooks, "Intelligence without representation," *Artificial Intelligence J.*, vol. 47, 1991, pp. 139-159.

[2] G.A. Carpenter and S. Grossberg, "*Pattern recognition by self-organizing neural networks*," MIT Press, MA, Cambridge, 1991.

[3] R.A. Baxter, "Supervised adaptive resonance networks," In *Proc. of the Conf. on Analysis of Neural Network Applications*, Fairfax, VA, 1991, ACM, pp. 123-137.

[4] G.A. Carpenter, S. Grossberg. N. Markuzon, J.H. Reynolds and D.B. Rosen, "Fuzzy ArtMap: neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. on Neural Networks,* 1992, vol. 3, pp. 698-713.

[5] Zhihua Zhou, Shifu Chen and Zhaoqian Chen, "FANNC: a fast adaptive neural network classifier*," Knowledge and Information Systems, vol.* 2(1), 2000, pp. 115-129.

[6] A.V. Gavrilov, V.V. Gubarev, K.-H. Jo and H.-H. Lee, "Hybrid neural-based control system for mobile robot," In *Proc. of 8th Korea-Russia Int. Symp. on Science and Technology KORUS-2004, vol. 1,* Tomsk, Russia, June, 2004, TPU, pp. 31-35.

[7] An-Min Zou, Zeng-Kuang Hou, Si-YaoFu and Min Tan, "Neural networks for mobile robots navigation: a survey," In *Proc. of Third Int. Symp. on Neural Networks ISNN-2006*, part 2, Chengdu, China, May, 2006, Springer, Advances in Neural Networks ISNN-2006, LNCS 3972, pp. 1218-1226.

[8] J. Hawkins and S. Blakeslee, "*On intelligence*," Times Books, 2004.

[9] D.E. Rumelhart, "Parallel distributed processing," In *Mcclelland J.L. (Eds.), Explorations in the Microstructure of Cognition,* vol.. I, II, MIT Press, MA, Cambridge, 1986.

[10] K. Fukushima, "Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.,* 1980, vol. 36, pp. 193-201.

[11] V. Petridis and V.G. Caburlasos, "Fuzzy lattice neural network (FLNN): a hybrid model for learning," *IEEE Trans. on neural networks*, vol. 9(5), 1998, pp. 877-890.

[12] Satoshi Suzuki and Hiroshi Ando, "A modular network scheme for unsupervised 3D object recognition," *Neurocomputing,* vol. 31, 2000, pp. 15-28.

[13] L. A. Torres-Méndez, J. C. Ruiz-Suárez, Luis E. Sucar, and G. Gómez, "Translation, rotation, and scale-invariant object recognition," *IEEE Tran. on Systems, Man and Cybernetics - Part C: Applications and Reviews*, vol. 30(1), 2000, pp.125-130.

[14] Jian Huang Lai, Pong C. Yuen and Guo Can Feng, "Face recognition using holistic Fourier invariant features," *Pattern Recognition,* vol. 34, 2001, pp. 95-109.

[15] Jung-Hua Wang, Jen-Da Rau, Wen-Jeng Liu, „Two-stage clustering via neural networks," *IEEE Trans. on Neural Networks,* vol. 14(3), 2003, pp. 606-615.

[16] Hahn-Ming Lee, Chih-Ming Chen, Yung-Feng Lu, "A self-organizing HCMAC neural-network classifier," *IEEE Trans. on Neural Networks*, vol. 14(1), 2003, pp. 15-27.

[17] A.V. Gavrilov, "Hybrid neural network based on models of multi-layer perceptron and adaptive resonance theory," In *Proc. of 9th Korean-Russian Int. Symp. on Science and Technology KORUS-2005,* Novosibirsk, Russia, June, 2005, NSTU, pp. 119-122.

[18] A.V. Gavrilov, Y.-K. Lee and S.-Y. Lee, "Hybrid neural network model based on multi-layer perceptron and adaptive resonance theory," In *Proc. of Third Int. Symp. on Neural Networks ISNN-2006*, part 1, Chengdu, China, May, 2006, Springer, Advances in Neural Networks ISNN-2006, LNCS 3971, pp. 707-713.

[19] T. Kohonen, "Self-organized formation of topologically correct feature maps," Biological Cybernetics, vol. 43(1), 1982, pp. 59-69.