Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 53(67), Fascicola 1, 2008

Genetic procedure for optimization of RBF neural network center positioning

Iulian C. Vizitiu, Petrică Ciotîrnae[1]

Abstract – According to RBF neural network theory, it is well known that the (recognition) performances of these architectures depend a lot by the positioning method of centers into input dataspace. Using the affiliation of genetic algorithms to the class of the global searching techniques, and their ability to offer very good results in solving of complex optimization problems, it is justified the attempt to select RBF neural network centers through a suitable genetic procedure.

Keywords: RBF neural network, genetic algorithm, pattern recognition

I. INTRODUCTION

According to [1], it is known that the *back-propagation* (BP) algorithm used in case of *multilayer perceptron* (MLP) training represents an optimization method by stochastic approximation type. Another procedure for a neural network design and training can be that to consider this problem as a curve approximation in R^n space and thus, to determine a n-D surface achieving the best matching with the input pattern set. Accordingly, the generalization capacity of this new neural network can be used with success for interpolation of the data from the testing set. Also, the hidden layer neurons will have the role to generate a function set for a suitable representation of each input pattern, this representation space being made from so-called *radial basis functions* (RBF) [2].
The mathematical support of RBF neural network theory is based on the *Covers* theorem alluding to the vectors (features) discrimination [3]: "*A complex pattern-classification problem cast in high-dimensional space nonlinearly is more likely to be linearly separable than in a low dimensional space*".

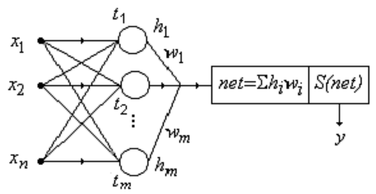


Fig. 1. n-D RBF neural network architecture

Consequently, an important step in design and training procedure of RBF neural networks consists in $\{t_i\}_{i=\overline{1,m}}$ center selection (see Fig.1). In the fundamental literature assigned to the theory of RBF neural networks, the *basic* strategies for center selection are: random positioning, supervised selection and clustering selection techniques [2], [3]. Also, based on these fundamental strategies, in literature a lot of suitable methods for RBF network center selection are mentioned [4], [5] ,[6], [7].
Generally speaking, the standard approches used for RBF network center selection lead to some important *disadvantages*, such us: difficulties due to network overfitting involving calculus time increasing, bad conditioning problems due to liniar dependence caused by center proximity etc. Consequently, all these disadvantages have as immediate effect a serious decreasing of the classification performances for a recognition system (ATR system) using RBF neural networks [8].
According to [9], an efficient and flexible solution for increasing the pattern recognition performances of RBF neural networks comparing to the case of standard center selection algorithms use can be *unsupervised competitive clustering* algorithm (UCC) proposed by Brown.
Taking into account the affiliation of *genetic algorithms* (GA) to the modern class of the global searching methods, the aim of this paper is to provide a design procedure for a specific GA which represents an improved and more robust alternative solution to the reference standard selection methods [9], [10].
Therefore, in the first part of the paper, a theoretical demonstration of GA design, as well as the basic properties of this new selection method is presented. In the last part of the paper, the experimental results that confirm the theoretical properties of the proposed GA are shown. Finally, some important conclusions and future research directions in this action field are also included.

II. GENETIC ALGORITHM DESIGN

The design procedure for the proposed GA starts with the solving of two basic problems assigned to genetic optimization [10]: *encoding* of the interesting real

[1] Military Technical Academy, Communications and Electronic Systems Department
George Cosbuc Avenue, no. 81-83, 050141, Bucharest, e-mail: vic@mta.ro, ciotirnae@mta.ro

problem in chromosomal terms and respectively, a suitable definition of the *fitness* function.

In order to optimize the parameters of RBF network functions $\{t_i, \sigma_i\}_{i=\overline{1,m}}$, the proposed GA must contain the following important two *steps*:

S1. if the training dataset has the form $\{\mathbf{X}_k, d_k\}_{k=\overline{1,P}}$, $\mathbf{X}_k \in \mathrm{R}^n$ and c is the number of the classes from input space, then using an *adaptive clustering* algorithm (with *zoom* effect, e.g. *k*-means or *ISODATA* techniques) the major tendencies from inside of each data cluster are determined. Therefore, based on this clustering method use, inside of each main data cluster from input space is bounded into $m_j \left(m = \sum_{j=1}^{c} m_j \right)$ new subclusters;

S2. the starting chromosome population is made using a random choice of m_j vectors $\mathbf{X}_k$ from each class (one vector for each bounded subcluster) and finally, a vector linear concatenation. Therefore, each chromosom will have assigned m vectors $\{t_i\}_{i=\overline{1,m}}$ which are extracted from the training dataset. Also, in order to provide a suitable representation, it was used a *real* encoding technique.

After the applying of RBF network center selection procedure, $\{\sigma_i\}_{i=\overline{1,m}}$ spread for each hidden neuron was calculated according to following equation:

$$\sigma_i^2 = \frac{1}{p_i} \sum_{x_i} (x_i - t_i)^T (x_i - t_i) \qquad (1),$$

where the used notations are consacreted [1].

Consequently, each output of RBF neural network can be written according to the following equation [2]:

$$y_k = \mathrm{S}\left(\sum_{i=1}^{m} h_i w_i \right) = \mathrm{S}\left(\sum_{i=1}^{m} w_i \exp\left(-\frac{d^2(\mathbf{X}_k, t_i)}{2\sigma_i^2} \right) \right), \quad k = \overline{1,P} \qquad (2),$$

where $\mathrm{S}(\cdot)$ is the transfer function for each neuron from output layer, by *purelin* type.

Because in this moment RBF setting parameters $\{t_i, \sigma_i\}_{i=\overline{1,m}}$ are known, the neural weigths to output layer $\{w_i\}_{i=\overline{1,m}}$ can be easily calculated using standard OLS algorithm [2].

The *fitness* function used for each chromosome evaluation is by RMS *error* type, and it was calculated according to the following equation:

$$E = \frac{1}{1 + \left[\frac{1}{P} \sum_{k=1}^{P} (y_k - d_k)^2 \right]^{0,5}} \qquad (3).$$

The *stopping criterion* for the proposed GA was represented by the exceeding of the maxim generation number (this number has a constant value) or when the goal error was reached.

The *parents selection* for the next generation was realized using an *elitist* method. To eliminate untimely convergence phenomenon, the *fitness* function was scaled acording to the following equation:

$$E_{\max}^{new} = k \cdot E_{\max}^{old}, \quad k \in [1.2, 2] \qquad (4).$$

The *crossover* supossed the use of two splitting points (randomly chosen), and each chromosome had attached a certain crossover probability with values into $[0.6, 0.95]$ range. In order to introduce new chromosomes inside of the current population, and to protect genetic algorithm against irreversible and accidental information failures generated by improper crossover operations, the *mutation* was used. The probability of mutation was chosen into $[0.001, 0.01]$ range.

Generally speaking, it is known that the solution given by genetic algorithm is coded under the form of the most performant chromosome that belongs to the last generation but in fact, nothing not garantees us that a more performant chromosome has not been obtained, for example, in the previous chromosomial generation. Consequently, using the analogy with the *Gallant* algorithm from neural network theory [9], at each chromosomial generation, the best chromosom from this population will be kept into *pocket* and thus, after a suitable decreased order technique, it will result indeed the best final solution.

More details regarding theoretical aspects treated in this section can be found in [9] and [10].

III. EXPERIMENTAL RESULTS

The main *objective* of experimental part of the paper is to demonstrate the superiority as performance level, comparing to standard UCC technique, of the proposed genetic center selection procedure.

The logical diagram used to generate the input video database for RBF neural network training and testing is presented in Fig.2.

As it can be seen from Fig.2, the video database used in this application was obtained from a (digital) photographical survey of *five* military aircraft models

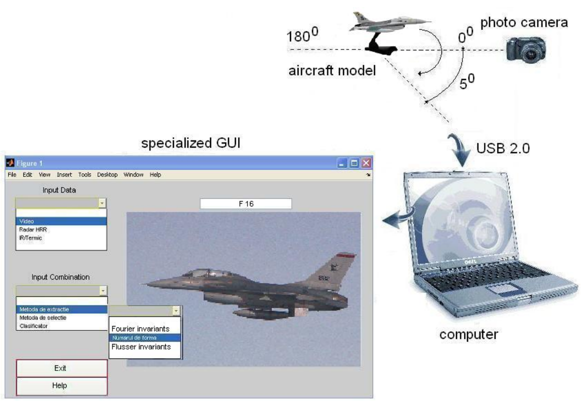
(F117, Mirage 2000, Mig 29, F16 and Tornado) scaled at 1:48 (see Fig.2a). The survey was taken using a 5^0 increment in the azimuthal plane, using a range of $\left[0^0, 180^0\right]$ justified by the geometric aircraft shape simmetry.

Each image from the input video database has a resolution of 520×160 pixels, in an uncompressed BMP format.

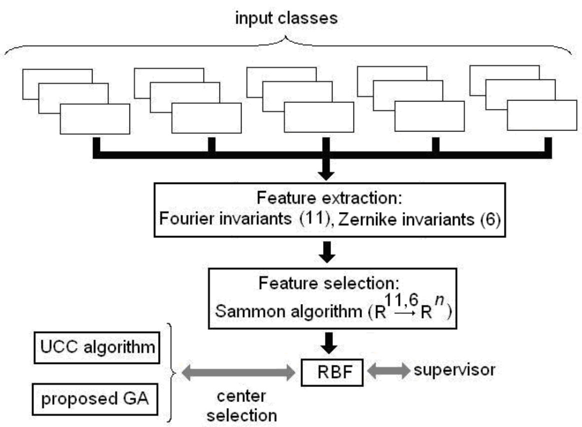
After the acquisition and preprocessing steps (see Fig.2b), a number of 37 video images/class is obtained. A number of 19 images were used for RBF neural classifier training while for testing 18 images were used. As feature extraction methods, Fourier and respectively, Zernike invariants were used, and as feature selection method, an *improved version* of Sammon nonlinear projection algorithm was used [9].

F 16 Mig 29

(a) Examples of military aircraft models used in database design



(b) Acquisition and preprocessing stage



(c) Center selection and training stage of RBF network

Fig.2. The testing procedure used in case of RBF network center selection

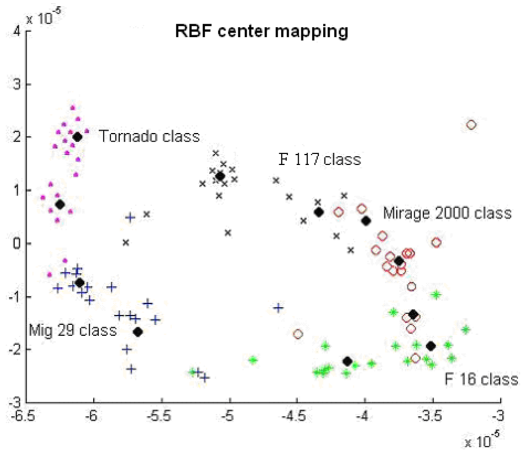
In order to implement and test the proposed optimization method, *Genetic optimization* toolbox (*GAOT*) from MATLAB™ 7.0 package was used, on a Pentium^{TM} processor at 2.4 GHz.

The experimental results obtained after RBF center selection method applying are indicated in Table 1. Also, after RBF center selection and training stage, it is important to quantify and analyze the following important *two* parameters: CR (*classification rate*) that represents, in (%), the ratio between the number of correct classified input patterns and the total number of patterns used for classification and respectively, CT (*convergence time*) that represents, in (s), the medium time needful to obtain the final RBF network classification solutions.

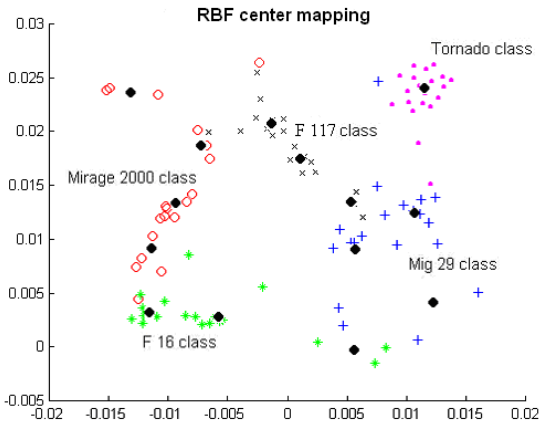
Table 1

RBF network center selection method	**Results**	**Running parameters**
Case of Fourier invariant use as feature extraction method (11)		
UCC algorithm	CR=91%, CT=0.67 s n=7	maxepochs=10^4 $\varepsilon=10^{-2}$ spread=0.9 m=8
Genetic algorithm	CR=94%, CT=0.54 s n=6	run time=175 s maxpop=75 maxstring=66 maxgen=100 p_i=0.8 p_m=0.05 e=6; $\varepsilon_0=10^{-2}$ c=5; P=185 m=11
Case of Zernike invariant use as feature extraction method (6)		
UCC algorithm	CR=93%, CT=0.7 s n=5	maxepochs=10^4 $\varepsilon=10^{-2}$ spread=0.85 m=9
Genetic algorithm	CR=96%, CT=0.6 s n=5	run time=184 s maxpop=75 maxstring=70 maxgen=100 p_i=0.8 p_m=0.05 e=6; $\varepsilon_0=10^{-2}$ c=5; P=185 m=14

In Fig.3a, at finish of GA running (using Fourier invariants), a 2D projection of RBF center mapping over input data space is shown. Also, in Fig.3b, the same graphical representation but using Zernike invariants is presented.



(a) Case of Fourier invarint use as feature extraction method



(b) Case of Zernike invarint use as feature extraction method

Fig.3. RBF network center 2D projection over input data space

As one can see in Fig.3, GA used for RBF network center selection leads to a very good center positioning over input data space (each significant data (sub)cluster has allocated at least a RBF center) even through it was used a minimal 2D representation. Therefore, it is expected to obtain also very good classification results (see Table 1).

More details regarding experimental aspects treated in this section of the paper can be found in [8].

V. CONCLUSION AND FUTURE WORK

The theoretical and experimental results presented in this paper leads to the following *remarks* concerning the proposed GA for RBF network center selection:

- applying of the proposed GA for RBF network center selection, leads to very good classification performances ($CR > 94\%$), and has as result a good increasing of CR, generally 3% more than standard UCC algorithm use. Also, using this genetic procedure, CT decreases, generally 20% less than reference algorithm (see results from Table 1);
- although the input database and recognition system structure analyzed in this paper are not identically with ones used in others references, however the increasing ratio of classification results is similar, comparing to other RBF network center selection methods [4], [5], [6], [7];
- generally speaking, the most important disadvantage of GA use for RBF network center selection is concerning the more increased computing resources needful to obtain the best individuals comparing to standard selection techniques. In our study case, this disadvantage is not very significant because the dimension of input database used for applications is medium.

As a conclusion, the design of GA for RBF network center selection is *feasible*, and this GANN system can be used in an efficient manner inside of real recognition system (e.g., ATR system).

In a future improvement, the proposed GA will contain in addition a *new* module intended for standard RBF network training algorithm (in this paper, the OLS algorithm was used) replacement. Accordingly, all steps belong to RBF network training process will become evolutionist (in fact, we will discuss about a *full* genetic procedure for RBF network design and training).

To increase the speed of genetic optimization process, another interesting point for a future development refers to a suitable *hardware* implementation of this new GA (e.g., based on FPGA technology use). Also, using this approach, the possibility to obtain the real dimension of genetic optimization process influence on the recognition (ATR) system performances will become relevant.

REFERENCES

[1] Duda R.O., Hart P.O., and Stork D.G., *Pattern Classification*, New York, USA: John Wiley&Sons, 2001

[2] Ghosh J., Nag A., *An overview of Radial Basis Function Networks*, Physica-Verlag, 2001

[3] Howlett R., Jain L.C., *RBF Networks: New Advances in Design*, Physica-Verlag, Heidelberg, 1st Edition, 2001

[4] Lacerda E.G., *Model Selection of RBF Networks via Genetic Algorithms*, Ph.D. Thesis, Pernambuco Federal University, Informatics Center, 2003

[5] Lee D.W., Lee J., "A Novel Three-Phase Algorithm for RBF Neural Network Center Selection", *Lecture Notes in Computer Science*, Springer Berlin, Heidelberg, vol.3173, 2004, pp.350-355

[6] Mao K.Z., "RBF neural network center selection based on Fisher ratio class separability measure", *Neural Networks, IEEE Transaction on,* vol.13, issue 5, 2002, pp.1211-1217

[7] Vizitiu I.C., "Genetic Algorithm for RBF Neural Network Center Selection", *Proc. of Int. Conf. WSEAS NN'08*, Sofia, Bulgary, 2008, pp.431-436

[8] Vizitiu I.C., "A Comparative Study concerning the Influence of RBF Neural Network Center Selection on Classification Performances", *Proc. of Int. Conf. ECAI*, Pitesti, Romania, 2007, pp.231-236

[9] Vizitiu I.C., *Genetic Algorithms used in Neural Network Optimization*, MTA Publishing House, 2005

[10] Cristea A., Zaharia M., *Genetic Algorithms and Neural Networks*, Romanian Academy Publishing House, Bucharest, 2002