A RBF/MLP Modular Neural Network for Microwave Device Modeling

Márcio G. Passos†, Paulo H. da F. Silva†† and Humberto C. C. Fernandes†

† Electrical Engineering Department, Federal University of Rio Grande do Norte, Natal - RN , 59072-970, Brazil
†† Federal Center of Technological Education of Paraíba, João Pessoa - PB, 58015-430, Brazil

Summary
This work presents a new Radial Basis Function/Multilayer Perceptron (RBF/MLP) modular structure, training with the efficient Resilient Backpropagation (Rprop) algorithm, that has been used for nonlinear device modeling in microwave band. The proposed modular configuration employs three or more nets, each one with a hidden layer of neurons. This method was proposed on the basis of the different characteristics of the two networks types: The MLP networks construct global approximations to nonlinear input-output mapping, consequently they are able to generalize in those regions of the input space where little or no training data is available. However, RBF networks use exponentially decaying localized nonlinearities to construct local approximations to nonlinear input-output mapping. Simulations through the proposed neural network models for microwave waveguide and patch antenna on PBG (Photonic Bandgap) structures and gave answers in excellent agreement with accurate results (measured or simulated) available in the literature.

Key words:
Neural networks, Data modeling, Computational methods

1. Introduction

Since the beginning of the 1990s, the artificial neural networks have been used as a flexible numerical tool, which are efficient and accurate for the RF/microwave device/circuit modeling. The neural models, which are trained by means of precise data (obtained through measurements or by electromagnetic simulation), are used in the design/optimization phase of devices and circuits, supplying fast and accurate answers.

In the CAD (Computer Aided-Design) applications related to microwave engineering and optical systems, the use of artificial neural networks as nonlinear models becomes very common, [1]. Recent publications, in the literature about this subject, indicate that: the use of previously established knowledge in the microwave area (as empirical models) in conjunction with the neural networks, results in a major reliability of the resulting hybrid model – with a major ability to learn nonlinear input-output mappings, as well as to generalize answers, when new values of the input are presented. Another important advantage is the data amount reduction necessary for the neural networks training used. Some modeling techniques have been proposed for the use with empirical models and neural networks, such as: Source Difference Method, [2], PKI (Prior Knowledge Input), [3], KBNN (Knowledge Based Neural Network), [4] and SM-ANN (Space Mapping Artificial Neural Network), [5].

A disadvantage in the hybrid models use is the need of an empirical model. When this becomes a limitation, for example, when a new component does not have an empirical model or an equivalent circuit, the EM-ANN (Electromagnetic - Artificial Neural Network), [1] conventional technique, is usually used. In this case, a simple neural network, MLP or RBF, is trained directly through EM/physics data which represent the functioning nonlinear model of the analyzed component.

The EM-ANN technique has been used in the neural network training as models of a microwave active and passive components variety, which presents a nonlinear behavior considered smooth, for example: transistors, discontinuity in microstrip lines and passive components, [1].

However, the EM-ANN technique presents some disadvantages which limit its application. For instance, as all the information is obtained through the ANN data training, a major amount of data is necessary to maintain the model accuracy. The increase of the training dataset size in a complex learning problem may overload a neural network, making its dimension and training difficult. On the other hand, even with a sufficient amount of training data, the reliability of the resulting neural models, when used for extrapolation, is not guaranteed, and, in many cases, it is very poor, [1].

The majority of the problems found in the EM-ANN technique use can be handled through the neural networks combination in modular structures which increase the training efficiency and the resulting neural model accuracy, [6], [7]. This concept is based on the principle divide and conquer in which a nonlinear modeling complex problem is divided in smaller problems, which are solved among the neural networks of the modular structure.

In this article an RBF/MLP modular structure is proposed through the combination of two expert RBF networks and an output MLP network. The development of models through the RBF/MLF modular structure is described in section 2. The applications of these neural models for microwaves devices with PBG periodic structures are described in section 3. A comparative study of the implemented models reliability through the MLP, RBF neural networks and RBF/MLP modular one is also included. Section 4 gathers the conclusions of this research.

2. Methodology through the RBF/MLP modular neural network

The proposed modular structure uses three feed forward neural networks, each one with a hidden neuron layer: two expert networks of the RBF kind and an output network of the MLP kind. Figure 1 presents a diagram in RBF/MLP modular structure blocks. This choice was motivated by the individual characteristics of the MLP and RBF networks, when used for the function approximation: the RBF network performs a local approach, serving as an expert network, since it grasps the models' nonlinearities; the MLP network performs a global approach and acts as an output network, since it favours the generalization capacity of the modular structure. The parameters of the model input, designed by 'initial value', 'final value' and 'intermediary value' are related to the interest region defined by the training data, Fig. 2. In order to receive additional information supplied by the pre-trained expert RBF networks, the output MLP network has two extra inputs, Fig. 1.

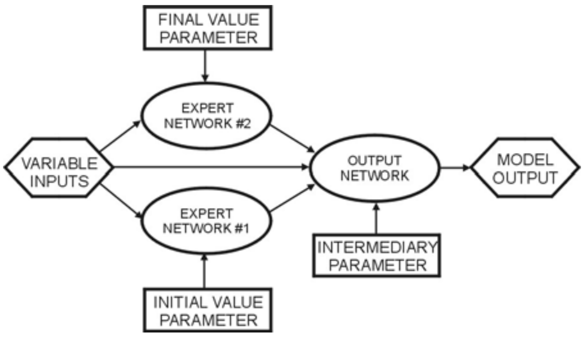

Fig. 1 The proposed modular network configuration.

The modeling problem mentioned is established by means of a normalized set of measured/simulated data, cited by $S = \{x(n), d(n)\}$, in that, $1 \le n \le$ N, and N is the total number of examples in the S training dataset. The x vector gathers the parameters of the model input (for instance, the gate length/width of a field effect transistor (FET); the length and the radius of a cylindrical antenna). The d desired answer describes the device EM/physics behavior under consideration (for instance, an FET drain current; the input impedance of a cylindrical antenna). The EM/physics theoretical relation between x and d is given by,

$$d = f(\boldsymbol{x}) \tag{1}$$

where, f represents the input-output mapping, which can be multidimensional and highly nonlinear. The aim is to develop a fast and accurate neural model for the f relation. The neural model is defined through the relation,

$$y = y(\boldsymbol{x},\boldsymbol{w}) \tag{2}$$

where, w represents the free parameters (or weights) of the neural network.

The use of the RBF/MLP modular structure enables the division of a modeling problem in smaller and easier problems to be solved. To describe this division, the interest region is taken into account defined through the training data for a hypothetical device, Fig. 2. The data referred to the 'initial value' and the 'final value' parameters are used in the training of #1 and #2 expert RBF networks, respectively; the training of the MLP output network is done with all the training data, including the 'intermediary values' available.

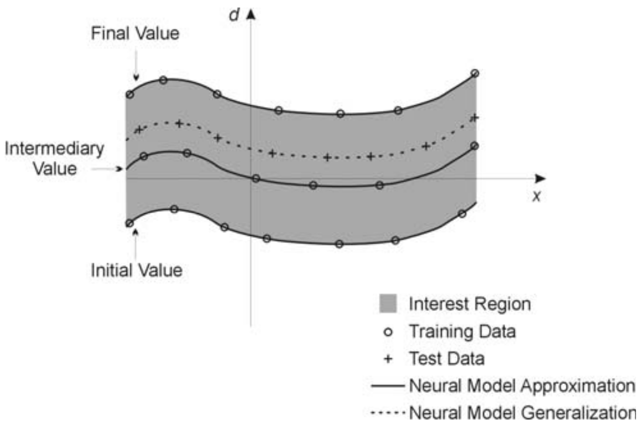

Fig. 2 Interest region defined by the training data.

In the MLP and RBF network supervised training with the backpropagation algorithm [1], the adjustment of the free parameters is carried out through the steepest descent method,

$$\boldsymbol{w}(n) = \boldsymbol{w}(n-1) - \eta \nabla E\big(\boldsymbol{w}(n-1)\big) \tag{3}$$

where, ∇ is the gradient operator; η is a training parameter, called learning rate, that controls the adjustments applied to the ANN's free parameters; and E is the square error, defined by,

$$E(n) = \frac{1}{2}e(n)^2 = \frac{1}{2}[d(n) - y(n)]^2 \tag{4}$$

in that, $e(n)$ is the instantaneous error between the desired answer and the neural network output. The training is carried out until the mean square error (MSE) reaches a minimum pre-established value. The MSE is a parameter that measures the training performance, being defined by,

$$\mathrm{MSE}(t) = \frac{1}{N}\sum_{n=1}^{N} E(n) \tag{5}$$

where, t is an index for the number of training epochs. An epoch is counted when all the training examples are presented to the neural network.

Due to the training slowness with the backpropagation algorithm, in this work, the use of the Rprop algorithm (using the standard training parameters) was chosen. The Rprop algorithm, proposed by Riedmiller and Braun [8], belongs to the algorithm family derived from backpropagation, which satisfies Jacobs' heuristics for the training acceleration, [9]. In an ANN training using the Rprop, just the gradient signs of the error function, Eq. (3), are taken into account. The negative influence elimination of the gradient amplitudes, as well as the use of adaptive and individual learning rates for each ANN free parameters, awards convergence speed and robustness as regards the choice of the training parameters of the Rprop algorithm, [8].

3. Models of Microwave Devices with PBG Periodic Structures

3.1 UC-PBG rectangular waveguide

The UC-PBG (uniplanar compact – photonic bandgap) rectangular waveguide, proposed in [10] designed for functioning in the X band, has lateral walls with UC-PBG cell periodic structures, Fig. 3(a), which in the resonant frequency act as magnetic surfaces, [10]. The electric field intensity as a function of the frequency and the position inside an UC-PBG rectangular waveguide was modeled through the RBF/MLP modular structure. The waveguide L measure is worth 22.86 mm and d measures 21.59 mm, Fig. 3(b). The UC-PBG metallic walls were built under a substrate of 0.635 mm thickness, with a dielectric constant, ε_r = 10.2.

In the neural models training for the UC-PBG waveguide, two input parameters were taken into consideration: the operation frequency, f, and the measurement position of the electric field, x. The measured values in the electric field make up the desired answers for the neural models. The training data were obtained through measurements presented in [10]. The information related to the RBF/MLP modular network training is presented in Table 1.

Figure 4 shows the approximations made by the expert RBF networks for the measured values of the electric field concerning the frequency, in the 'initial value' positions (x = 0.25) and of 'final value' (x = l) of the UC-PBG waveguide. Figure 5 presents the answers of the RBF/MLP modular model developed. A good agreement between this model's answers and the measured data was verified, with an excellent approximation capacity and generalization around 9.6 – 11 GHz.

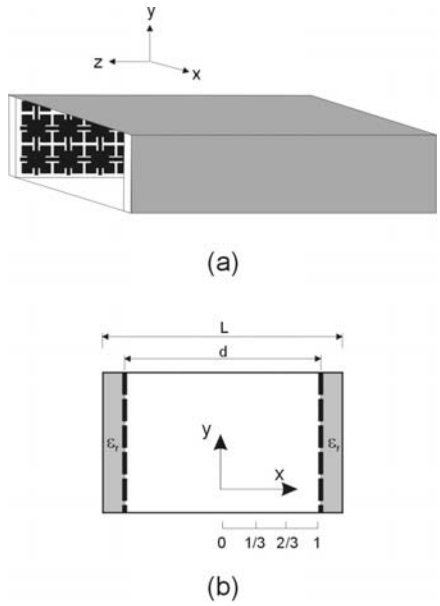

Fig. 3 UC-PBG Rectangular waveguide: (a) General view; (b) Transversal section.

Table 1: Information related to the RBF/MLP modular neural training for the UC-PBG waveguide.

Neural Network	Expert 1 RBF	Expert 2 RBF	Output MLP
Input parameter:	x = 0.25	x = 1	x=[0.25 0.5 1]
# hidden neurons:	10	10	10
# training data:	10	10	40
final MSE:	2.06E-6	1.13E-5	9.32E-5
# training epochs:	10000	10000	10000

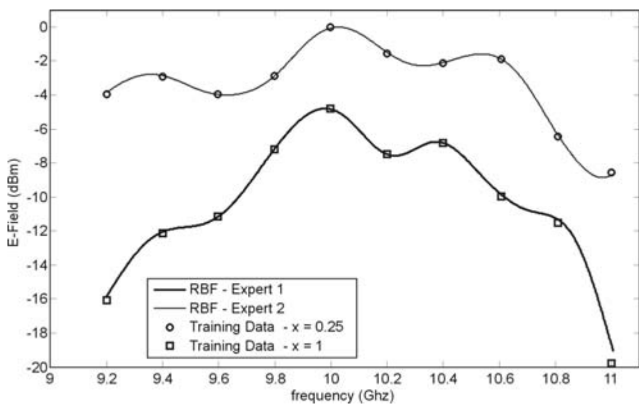

Fig. 4 Answers from the expert RBF networks for x = 0.25 and x = 1.

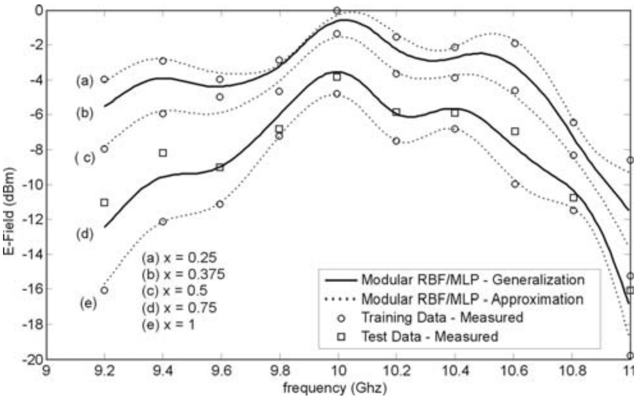

Fig. 5 Model answers through the RBF/MLP modular structure.

Aiming at verifying the reliability of the implemented models for the UC-PBG waveguide, through the MLP, RBF neural networks, and RBF/MLP modular ones, the number of hidden neurons in the single MLP and RBF networks and the MLP output network of the modular structure was noticed. This influences the generalization capacity of the resulting neural models. For each neural model, the mean square error was computed for a new test dataset (which was not used during the neural network training), corresponding to the position *x* = 0.75. Figure 6 presents the obtained results. It is noticed that, for the same number of hidden neurons, the RBF/MLP modular network has a major generalization capacity, showing an MSE smaller than the ones obtained with the use of the single MLP or RBF networks.

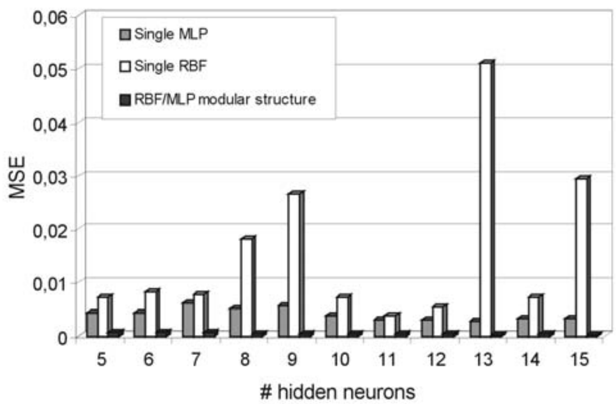

Fig. 6 Generalization capacity test of the MLP, RBF neural models and RBF/MLP modular for the UC-PBG waveguide regarding the number of hidden neurons.

3.2 Patch antenna with PBG substrate

One of the biggest disadvantages of the patch antennas is the loss due to the surface waves. The use of a PBG substrate enables the reduction of these losses. In this example, an RBF/MLP modular structure was used to shape the return losses in patch antennas with PBG substrate, through the mapping of the scattering parameter $|S_{11}|$ in function of the PBG substrate height and frequency.

Figure 7 illustrates a patch antenna with PBG substrate, whose analysis was made through the FDTD (Finite Difference Time Domain) method, [11]. This method is used to directly solve Maxwell's equations in time domain. Although it is a rigorous electromagnetic method, the FDTD presents a high computational cost, that, in general, its use in CAD applications becomes prohibitive.

As indicated in Fig. 7, the PBG substrate is formed by dielectric blocks, ε_r = 10.2; for the substrate remain, ε_r = 2.2. The rectangular patch has dimensions of 12.45 mm x 16 mm; the feeding line presents a width of 2.46 mm and a length of 8 mm, [11].

In the RBF/MLP modular model elaboration for the PBG substrate patch, two input parameters were taken into consideration: the frequency, *f*, in the 2.5-20 GHz band, and the PBG substrate height, *h*, between 0.794 mm and 1.588 mm. The training data were obtained by means of electromagnetic simulation with use of the FDTD method, [11]. Table 2 presents the relative information to the modular structure RBF/MLP training.

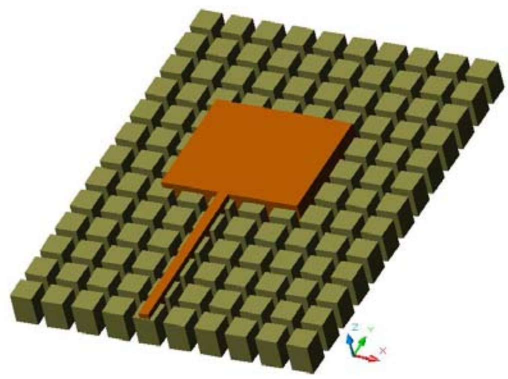

Fig. 7 Patch antenna with PBG substrate.

Table 2: Information related to the RBF/MLP modular neural training for the patch antenna with PBG substrate.

Neural Network	Expert 1 RBF	Expert 2 RBF	Output MLP
Input parameter:	*h* = 0.794	*h* = 1.588	*h* =[0.794 0.953 1.588]
# hidden neurons:	15	15	12
# training data:	47	46	139
final MSE:	1.87E-4	1.47E-4	2.06E-4
# training epochs:	10000	10000	30000

The approximations made by the expert RBF networks for the simulation results through FDTD method, referring to the heights of PBG substrate, *h* = 0.794 mm and *h* = 1.588 mm, are illustrated in Fig. 8.

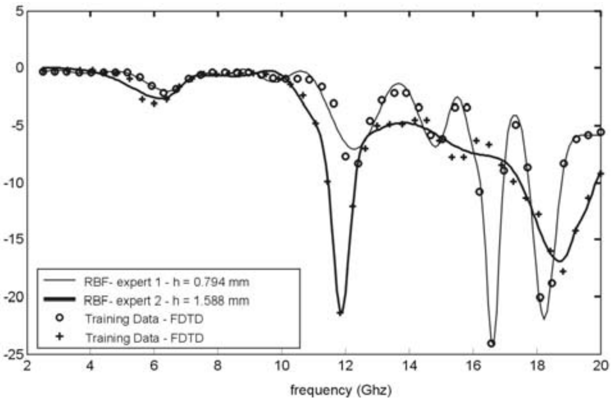

Fig. 8 Answers from the expert RBF networks

Fig. 9 presents the approximation performed by the RBF/MLP modular structure for the corresponding training data *h* = 0.953 mm. The results demonstrate the excellent capacity of the neural model approximation, even for a highly nonlinear mapping.

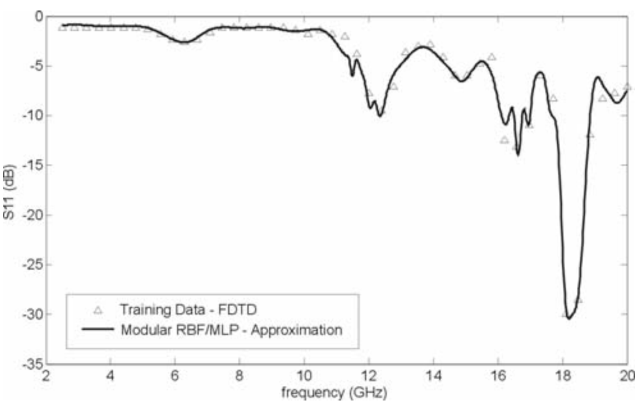

Fig. 9. Model answer through the RBF/MLP modular structure for *h* = 0.953 mm.

Figure 10 presents the answer of the RBF/MLP modular structure for the test data, correspondent to *h* = 1.429 mm. The good agreement between the neural model answers and the simulation results through FDTD method, demonstrates a good generalization capacity of the model through the RBF/MLP modular structure, mainly in the 6-14 GHz band.

In order to verify the trustworthiness of the models implemented for the patch with PBG substrate, through the MLP, RBF neural networks and RBF/MLP modular, it was verified as the number of hidden neurons of single MLP and RBF networks, and of the output MLP network of modular structure influences the generalization capacity of the resultant neural models. For each neural model, the MSE test was computed for the height of PBG substrate, *h* = 1.429 mm. Fig. 11 presents the obtained results. In relation to the models through MLP and RBF networks, it is verified that RBF/MLP modular model learning with consistency and presents a major generalization capacity, practically independent of the number of hidden neurons used.

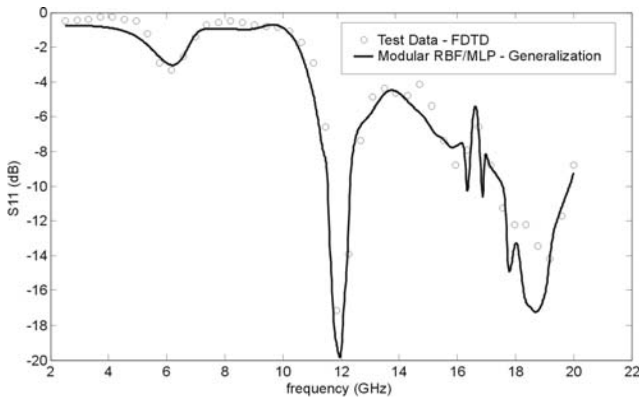

Fig. 10. Model answer through the RBF/MLP modular structure for *h* = 1.429 mm.

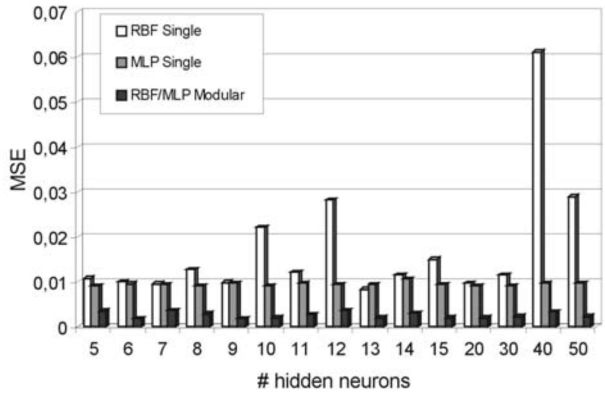

Fig. 11 Generalization capacity test of the MLP, RBF neural models and RBF/MLP modular for the patch antenna with PBG substrate regarding the number of hidden neurons.

4. Conclusions

In this paper a new RBF/MLP modular structure of neural networks, trained through the Rprop efficient algorithm, and developed specially for use in modeling applications, was proposed. In particular, an UC-PBG rectangular waveguide and a patch antenna with PBG substrate were used.

The RBF/MLP modular structure modules were organized in order to take advantage of the local and global approximation characteristics presented by the RBF and MLP neural networks, respectively. This kind of organization in conjunction with the modeling problem division, makes easier the expert RBF networks training and the output MLP network of modular structure. The neural models simulation results implemented, indicate a major learning consistency, or generalization, and a major reliability of the models developed through the RBF/MLP modular structure in relation to the ones developed through MLP or RBF single structures. Besides, the RBF/MLP structure, directly trained by means of measured/simulated data through the EM-ANN technique, is very flexible, and it still can be applied as models, mainly when new components/technologies for microwaves circuits arise.

Acknowledgments

This work was supported in part by CNPq and CAPES.

References

[1] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*, 1st ed. Norwood, MA: Artech House, Inc., 2000.

[2] P. M. Watson and K. C. Gupta, "EM-ANN models for microstrip vias and interconnects in multilayer circuits", in *IEEE Trans. Microwave Theory and Techniques*, 1996, vol. MTT-44, pp. 2495-2503.

[3] P. M. Watson, K. C. Gupta and R. L. Mahajan, "Development of knowledge based artificial neural networks models for microwave components", in *IEEE MTT-S Int. Microwave Symp. Dig.*, Baltimore, 1998, pp. 9-12.

[4] F.Wang and Q. J. Zhang, "Knowledge based neuromodels for microwave design", in *IEEE Trans. Microwave Theory and Techniques*, 1997, vol. MTT-45, pp. 2333-2343.

[5] J. E. Rayas-Sanchez, "Neural space mapping methods for modeling and design of microwave circuits", Ph. D. dissertation, McMaster University, 2001.

[6] S. Hashem, and B.Schmeiser, "Improving model accuracy using optimal linear combinations of trained neural networks", *IEEE Trans. Neural Networks*, Vol. 6, 1995, pp. 792-794.

[7] Lendaris, G. G., A. Rest, and T. R. Misley, "Improving ANN generalization using a priori knowledge to pre-structure ANNs", *Proc. IEEE Intl. Conf. Neural Networks*, Houston, TX, june 1997, pp. 248-253.

[8] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The Rprop algorithm", in *Proc. IEEE International Conference on Neural Networks*, vol. 1, 1993, pp. 586-591.

[9] R. A. Jacobs, "Increase rate of convergence through learning rate adaptation", *Neural Networks*, Vol. 1, 1988, pp. 295-307.

[10] F. Yang, K. Ma, Y. Quian and T. Itoh, "A Uniplanar Compact Photonic Band-Gap (UC-PBG) structure and its applications for microwave circuits", in *IEEE Trans. Microwave Theory and Techniques*, 1999, pp. 1509-1514.

[11] J. F. Almeida, C. L. S. Sobrinho and R. O. Santos, "Analysis by FDTD method of a microstrip antenna with PBG considering the substrate thickness variation", in *Proc. International Conference on Applied Electromagnetics and Communications*, 2003, pp. 344-347.

Márcio Galdino Passos received the B.S. in Telecommunications Technology from the Federal Center for Technological Education – CEFET-PB, Brazil, in 2004, and is currently working toward the M.S. degree in electrical engineering at the Federal University of Rio Grande do Norte-UFRN, Brazil. His research interests are computational methods and neural networks in microwave/optical engineering.

Paulo Henrique da Fonseca Silva received the B.S.E.E. degree from the Federal University of Rio Grande do Norte, RN, Brazil, in 1996, and the M.S.E.E. degree from the Federal University of Rio Grande do Norte, RN, Brazil, in 1998. He received the doctoral degree in electrical engineering from the Federal University of Campina Grande, PB, Brazil, in 2002. He is currently teaching at the Federal Center of Technological Education of Paraíba, João Pessoa, PB, Brazil. His interest research areas are microwave device analysis, neural networks, antennas, and propagation.

Humberto César Chaves Fernandes was born in Martins-RN, Brazil. He received with laude the B.S. in Electrical Engineering from the Federal University of Rio Grande do Norte-UFRN, Brazil in 1977, the M.S. (1980), PhD (1984) degrees and Postdoctoral program (1986) from the State University of Campinas-UNICAMP, Brazil. His current research interests are microwave, millimeter waves, smart antennas array, superconductivity, semiconductor, neural networks, electromagnetic, photonics, dynamic methods and applications. Prof. Fernandes has more then three hundred published works. Since 1978 he is at the Electrical Engineering Department from the UFRN, where he is a Senior Researcher and Professor. Prof. Fernandes is member of the SBrT, IEEE CONSOC, SBMO and SBPMat.