Proceedings of ISAP2007, Niigata, Japan

Design of Waveguide Matched Load Based on Multilayer Perceptron Neural Network

#TIAN YuBo, ZHANG XiaoQiu, and ZHU RenJie

School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang Jiangsu 212003, P. R. China, Email: yuboe@sohu.com

Abstract: Rectangular waveguide terminal matched load is designed based on multilayer perceptron neural network (MLPNN) that is trained by error back propagation (BP) method. One of improvements is that the structure of MLPNN can be adjusted automatically according to complicacy of modeled system. Simultaneously, objective function of MLPNN is modified based on information entropy theory. Training and testing sample are sorted randomly to make them more effective. As a result, H-plane T-kind terminal matched load of rectangular waveguide is designed successfully.

Key Words: Neural network, waveguide, matched load, entropy

1. Introduction

Artificial Neural Networks (ANNs) have many advantages, such as good learning ability, less memory demand, suitable generalization, fast real-time operating, simple and convenient to utilize, and so on. Therefore, it has become a research hotspot in past few years. Recently, ANNs have been applied to microwave engineering [1]. Reference [2] modeled microwave device using ANN, reference [3] calculated the resonance frequency of resonance cavity based on ANN, reference [4] applied ANN to measurement of RF and microwave, and reference [5] presented the synthesis of a printed dipole antenna in terms of ANN combined with FDTD. Multilayer perceptron neural network (MLPNN) with error back propagation (BP) learning method is often used in ANNs. However, it has three main disadvantages, namely slow convergent speed, existing local minimum and poor generalization. To overcome them, we propose some reformative methods. One of them is that the network structure can be adjusted automatically according to complexity of simulated system. Another is about error function. Least mean square (LMS) that is often adopted as error function is modified by adding entropy part according to Shannon information theory. Simultaneously, there are some other improvements, such as linear scale for input and output sample to make the network more steady and more generalization, random sorting of input and output sample in every training cycle of network to make the training more effective, and moment term added in training process to make the network more steady and easy to convergence. All of these are included in part Ⅱ. In part Ⅲ, terminal small matched load of rectangular waveguide is modeled and optimized based on the proposed MLPNN. Conclusions are given in section Ⅳ.

2. Multilayer perceptron neural network

Multilayer perceptron neural network that is used extensively is the one of artificial neural networks. Usually, MLPNN includes input layer, output layer and several hidden layer. Suppose that the network has p inputs, q outputs, and activation function of hidden layer adopts continuous function (such as Sigmoid function). The relationship between input and output may be regarded as nonlinear mapping from p-dimension Euclidian space to q-dimension Euclidian space. This kind of mapping may arbitrarily approach any continuous function, and it is a general function approximator. Usually, training method of MLPNN is error back propagation. Unfortunately, the traditional BP algorithm has

ISBN: 978-4-88552-223-9 C3055©IEICE

several serious disadvantages, such as slow convergence speed, existing local minimum and poor generalization and so forth. In this paper, we do some works to improve the network's performance as follows:

(1) MLPNN structure can be adjusted according to complex degree of simulated system. The adjustment includes number of hidden layers and number of neurons of every hidden layer. For MLPNN, the number of hidden layers depends on complexity of nonlinear mapping. Although it has been proved that one layer network is enough to accurately approach any nonlinear function, it is time consuming to find the optimization. The training efficiency of MLPNN is high if the network layers are increased suitably. When number of hidden layer is assured, number of neurons of every hidden layer is important to the training of MLPNN. In the same way, it need less neurons if the mapping is sample, and large neurons versus complex mapping. However, too more hidden layer and neurons are not suitable because it may cause network over-fitting and affect network generalization.

(2) Felicitously design objective function of MLPNN. Generally, the objection function of MLPNN is least mean square error. However, this kind of objective function may cause over-fitting of network. An adjusted objective function including entropy part is adopted in this paper to overcome the problem:

$$ObjFun = g \cdot \frac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} \left(o_i - y_i\right)^2 + \left(1-g\right) \cdot \frac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} \left\{ o_i \log \frac{o_i}{y_i} + \left(1-o_i\right) \log \frac{1-o_i}{1-y_i} \right\} \tag{1}$$

Where $ObjFun$ means objective function, N_{sample} number of sample, o_i output of sample, y_i output of MLPNN, and g filter parameter that can vary automatically with training process. The first part of expression (1) presents the mean square error of network, and the second part is relative entropy of o_i and y_i that shows their Kullback-Leibler distance.

(3) In every training cycle, the training and testing sample are sorted randomly, which may use them fully and make the training more effective.

(4) To make network more universally, the input and output sample may be scaled linearly. The scaling function and inverse scaling function are giver by:

$$\tilde{x} = \tilde{x}_{\min} + \frac{x - x_{\min}}{x_{\max} - x_{\min}} \left(\tilde{x}_{\max} - \tilde{x}_{\min}\right) \tag{2}$$

$$x = x_{\min} + \frac{\tilde{x} - \tilde{x}_{\min}}{\tilde{x}_{\max} - \tilde{x}_{\min}} \left(x_{\max} - x_{\min}\right) \tag{3}$$

The meaning of every symbol is in reference [7] and omitted here.

(5) Other improving techniques are also used, which include adding momentum item that speeds up the convergence, suitable initial weights that accelerate network constringency, and adding noise with Gauss distribution into input sample that mends the network generalization.

For clearness, Fig.1 shows the flowchart of the MLPNN.

3. Application of MLPNN

Usually, antenna array needs many terminal matched loads. Simultaneously, there is a contradiction between their absorbing performance and their size, weight in working bandwidth. Therefore, terminal matched load has to be optimized. The problem's essential is optimal design of small terminal matched load of rectangular waveguide. For improving the absorbing performance and reducing size and weight, matched load may adopt various structure forms, such as multilayer and multi-zone, H-plane T-kind, E-plane T-kind etc. According to adopted structure, different computing methods may be used, for example FEM, Genetic Algorithm (GA) and so forth. It is noticed that S parameter or equivalent current parameter of the system are decided by structure forms and its sizes, and there are nonlinear mapping between them. Furthermore, the nonlinear mapping can be approached well by ANNs. Therefore adopting ANNs to design matched load may not only ensure design precision but also speed up the design. In this paper, H-plane T-kind matched load of BJ-100 standard rectangle waveguide (see Fig.2) is designed based on MLPNN above mentioned. The MLPNN is with 4 input parameters respectively denoted by D_1 and D_2 (varying range is from 1mm to 3mm founded on experience), W_1 and W_2 (central symmetry, upper layer varying range is from 4mm to 12mm, low layer is from 8mm to 22.86mm founded on experience), and 1 output parameter denoted by $|S_{11}|$ (amplitude of reflection coefficient). The central frequency is 9.6GHz, and 2# material with $e_r = 7.043 - j0.451, m_r = 1.153 - j0.342$ is selected. Frequency and material are not included in MLPNN for reducing the network structure.

Some training data must be chosen to train the MLPNN. Furthermore, the training data have to be representative. Modeling and training the network is complex if the training data are too much. Simultaneously, it may cause the network over-fitting. However, the network comprehensiveness may be bad if the training data are few because of their non-completeness. The principle of Central Composite Design and the method of DoE are applied to determine the input vectors to ensure their whole ranges. The capacity of training sample is 81 and the testing sample is another 81 to check the trained MLPNN. Network training and testing sample may be gotten by experiment measurement or by electromagnetic simulators. For simpleness, FDTD method is used to obtain them.

To the designed MLPNN in this paper, its initial hidden layer is 2, and initial number of per layer is also 2. When training finished, the structure of network is 4×8×8×1. Figure 3(*a*) shows the MLPNN training result vs. computing result by FDTD, and ObjFun = 0.000021. Figure 3(*b*) shows the MLPNN testing result vs. computing result by FDTD, and ObjFun = 0. 000153. From Fig.3, we can conclude that training precision of network is high, and test precision may satisfy engineering demand. Fig.4 shows the relationship of ObjFun vs. epoch, and the curve changes basically according to the rule of exponential drop, which means the network is stable and convergent. In order to finding the optimal structure of matched load of rectangular waveguide, Genetic Algorithm (GA) is used based on the trained MLPNN, and the optimal result is W_1=22.86mm, D_1=2.06mm, W_2=6.0mm, D_2=1.017mm. In this situation, $|S_{11}|$ may exceed -26dB from 9.4 GHz to 9.8 GHz.

4. Conclusion

A reformative multilayer perceptron neural network was developed to model and optimize H-plane T-kind terminal small matched load of rectangular waveguide. The improvements include network structure adjusted automatically, modified error function, linear scaling and random sorting of input and output sample, adding moment term and so forth. From the computing results, we can conclude that the proposed network has some advantages, such as simple, robust, good generalization and so on. The methodology may be extended, in principle, to other similar microwave engineering analyses and designs.

References

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

[2] P. Waston, K. C. Gupta, R. L. Mabajan, “Developmetn of knowledge based artificial neural network models for microwave components”, IEEE MTT-S Digest, pp 9-12, 1998.

[3] R. K. Mishra, A. Patnaik, “Neurospectral computation for complex resonant frequency of microwave resonators”, IEEE Microwave and Guided Wave Letters, Vol. 9, No. 9, pp 351-353, 1999.

[4] J. A. Jargon, K. C. Gupta, D. C. DeGroot, “Applications of artificial neural networks to RF and microwave measurements”, Int. J. RF Microwave Computer-Aided Eng., No. 12, pp 3-24, 2002.

[5] H. J. Delgado, M. H. Thursby, “A novel neural network combined with FDTD for the synthesis of a printed dipole antenna”, IEEE Trans. AP, Vol. 53, No. 7, pp 2231-2236,2005.
[6] D. E. Rumelhart, G. E. Hinton, R. J. Williams, “Learning representations by back-propagation errors”, Nature, No. 323, pp 533-536, 1986.
[7] Q-J Zhang, K. C. Gupta, V. K. Devabhaktuni, “Artificial neural networks for RF and microwave design – from theory to practice”, IEEE Trans. MTT, Vol. 51, No. 4, pp 1339-1350, 2003.

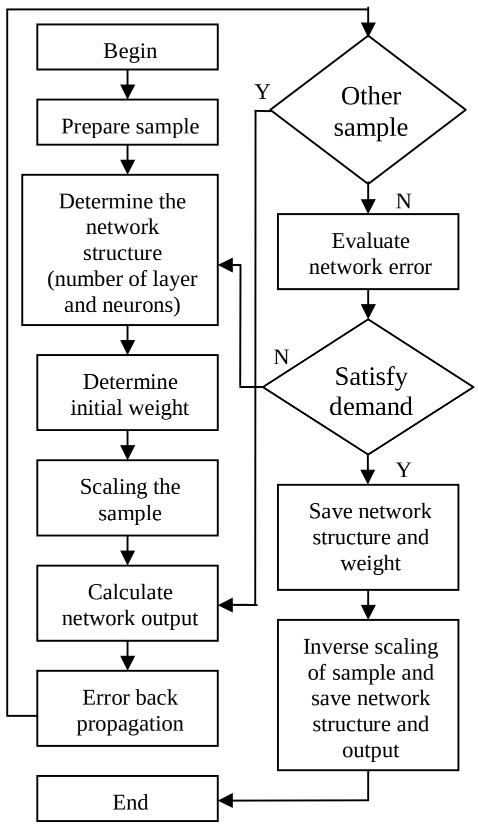
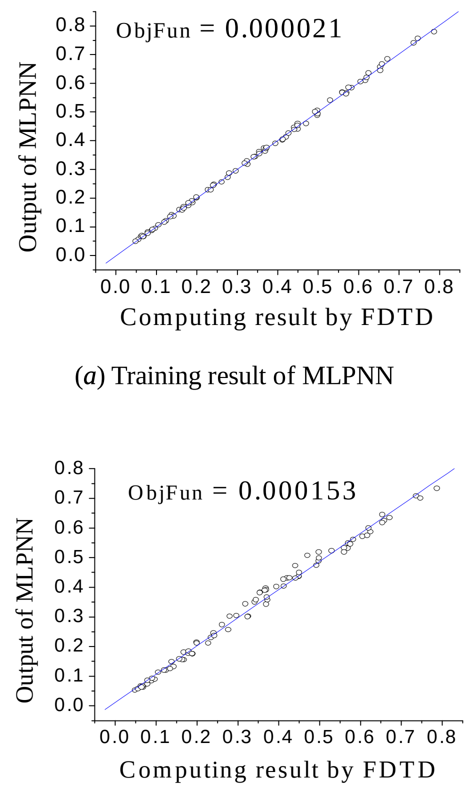


Fig. 1 Flow chart of MLPNN



(*a*) Training result of MLPNN

(*b*) Testing result of MLPNN

Fig. 3 Comparison of computing results by MLPNN model and FDTD

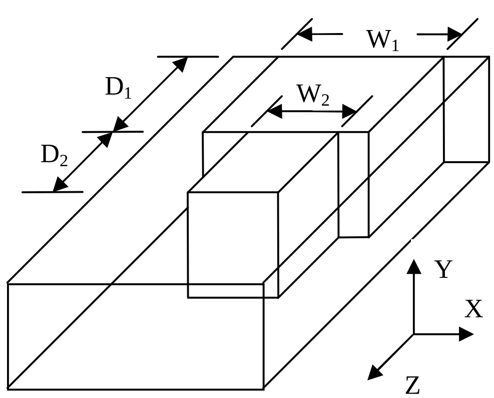


Fig. 2 H-plane T-kind matched load

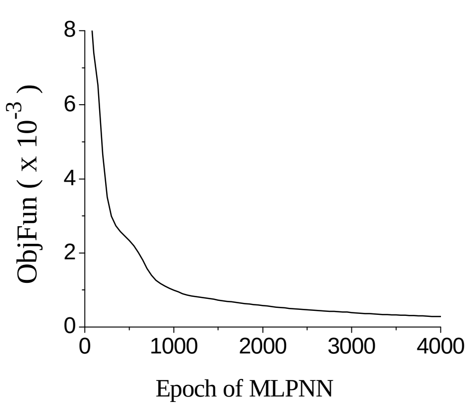


Fig. 4 Relationship of ObjFun vs. epoch