MICROSTRIP CIRCUIT DESIGN USING NEURAL NETWORKS

Tzyy-Sheng Horng and Chua-Chin Wang

Electrical Engineering Department, National Sun Yat-Sen University, Taiwan, R.O.C.

Nicolaos G. Alexopoulos

Electrical Engineering Department, University of California, Los Angeles, USA

Abstract

A new technique using neural networks to efficiently design microstrip circuits is presented. In our proposed method, a full-wave analysis is employed to rigorously characterize a microstrip circuit, which results in a finite set of pairs of input and output parameter vectors. The neurons, arranged as a three-layer network, are used to learn the mappings from input to output and then give accurate approximations for the output vectors at any arbitrary input. It is emphasized that a three-layer neural network is capable of performing any mapping if the right connections among the neurons can be made. A real example on the microstrip corporate feed design is given to illustrate the potential power of this technique.

I. Introduction

Neural networks have recently drawn significant attention as powerful tools in artificial intelligence (AI) research [1]-[4]. Their structural and behavioral resemblance to systems of biological neurons has also proved useful in a wide variety of engineering applications [5],[6]. In this paper, we wish to be the first to exploit the advantages using neural networks for microstrip circuit designs. In the past, several full-wave numerical methods including the space-domain and spectral-domain EFIE [7]-[12], MPIE [13],[14] and FDTD [15],[16] have been employed to effectively investigate various microstrip components and simple circuits. However, each method takes tremendous computation efforts which still can not make a practical circuit design feasible within a reasonable period of time. Therefore, an extensive literature was devoted to approximating an unknown mapping related to a circuit performance, in between or beyond the limited sampling points, by the curve-fitting techniques. This allows us to quickly calculate the output values at any arbitrary input points. In the curve-fitting techniques, the specific form of a function to be fitted to data is first chosen and then fitting according to some error criterion (such as the least mean of squared errors) is carried out. The functional form should be sufficiently general so as to approximate large classes of mappings which might arise in practice. The commonly used fitting functions include polynomials, rational functions and trigonometric functions.

A primary advantage of neural networks over the curve-fitting techniques is that the neural networks have more general functional forms. Kolmogorov's Mapping Neural Network Existence Theorem further proves that a three-layer neural network can exactly implement any continuous mapping [3]. This result gives hope that, with the proper connections among neurons, the network is itself able to approximate any mapping of practical interest to any desired degree of accuracy. Another advantage is neurocomputing's parallel distributed processing (PDP) ability to construct an unknown mapping from the tabulated output values at random input points [4]. The PDP models assume that information processing takes place through the interactions of a large number of simple processing neurons, each sending excitatory and inhibitory signals to other neurons in the learning procedure. This corresponds to a numerical recursive procedure for computing the weight changes of connectivity between any two neurons. Finally, a mapping can be in general represented by several weight matrices that constitute the pattern of connectivity from input to output.

In summary, many examples have been gathered to show that the function approximations using neural networks are usually better than those provided by the curve-fitting techniques which often exhibit some unwanted artifacts such as excessive polynomial-type humps or Fourier series-type overshoots and ringing. This difference is particularly evident in high-dimensional spaces (input dimensions greater than 3) [3].

II. Design of a Microstrip Corporate Feed

A design example for a microstrip corporate feed is given. The aim of the design is to ingeniously shape each junction in the corporate feed to acquire a tapered and in phase output current distribution from which a Dolph-Chebyshev array sum pattern may be synthesized. Fig. 1 shows a symmetric microstrip corporate feed with five output branch

CH3277-1/93/0000-0413$01.00 © 1993 IEEE

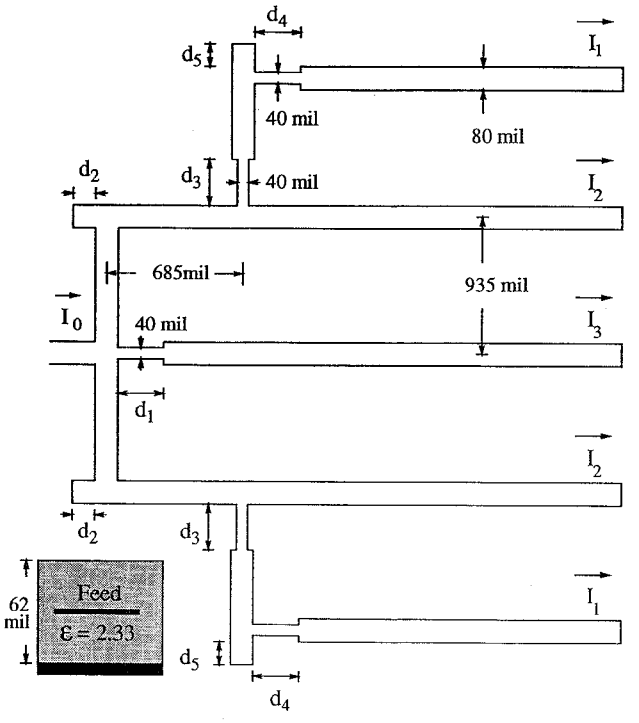


Figure 1: A prototype of microstrip corporate feed with five branch lines in the output ports.

lines. This corporate feed is embedded in the middle of a duroid substrate with permittivity equal to 2.33 and thickness equal to 62 mil. The operating frequency is 8.5 GHz. The microstriplines in the input ports, output ports, and between any two junctions are 50 Ω transmission lines. The corresponding width of microstriplines is equal to 80 mil and a guided wavelength is equal to 935 mil. In this corporate feed, we proposed five adjustable distance parameters $d_n, n = 1, \cdots, 5$, to meet the requirement for the specified array performance. From array theory, we know that the array sum patterns can be determined by two output current ratios, $\frac{I_1}{I_3}$ and $\frac{I_2}{I_3}$. These two ratios are complex values which include two magnitudes (denoted by t_1 and t_2) and two phases (denoted by t_3 and t_4). We can define a mapping $F : R_5 \rightarrow R_4$ by

$$F(d_1, d_2, d_3, d_4, d_5) = (t_1, t_2, t_3, t_4). \quad (1)$$

Our goal is to find a representation for this mapping using neural networks.

III. Full-Wave Junction Analysis

The first step to reach the goal is to establish a set of samples from which the neural networks may learn how output will respond to input. In this paper, we characterize each junction in the corporate feed by the scattering parameters at discrete values of certain distance parameters. Network theory is then applied to integrate all the junctions' scattering matrices and compute the output current distribution. The approach to find the scattering parameters is governed by the spectral-domain EFIE [10]-[12]. As an example, Fig. 2(a) shows the geometry of a shaped cross junction which is the first junction to divide the input currents. In the moment method procedure, the electric surface currents on the microstrips are effectively expanded by a combination of the rectangular subdomain functions and the semi-infinite traveling wave functions. The use of this expansion mechanism is also illustrated in Fig. 2(a) where each pair of adjacent rectangles represents a rectangular subdomain function. The dashed lines in the figure locate the reference planes. The semi-infinite traveling wave functions extend from these reference planes to expand in part the currents along the microstriplines. It is noted that we can easily lengthen or shorten the distance d_1 by adding or removing some rectangles without changing the entire numerical scheme for the characterization of this junction. As a result, magnitudes and phases of the scattering parameters at uniformly sampled values of d_1 are shown by the dark circles in Fig. 2(b) and Fig. 2(c) respectively.

IV. Framework of the Neural Network

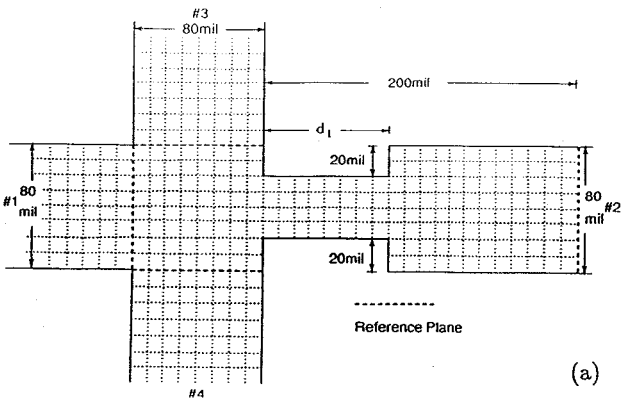
Fig. 3 shows a three-layer neural network. The patterns of connectivity among the neurons are represented by two weight matrices, $[w^a_{ji}]$ and $[w^b_{kj}]$. The weight is a positive number for exciting the connected neuron, and a negative number for inhibiting the connected neuron. Its absolute value specifies the strength of the connection. For the neurons in the last two layers, a continuous, nonlinear activation function is needed to bound the output signal and make sure that the network can effectively learn to approach a mapping. Here we choose a commonly used activation function in biological experiments as

$$x^{out} = f(x^{in}) = \frac{1}{1 + e^{-x^{in}}}. \quad (2)$$

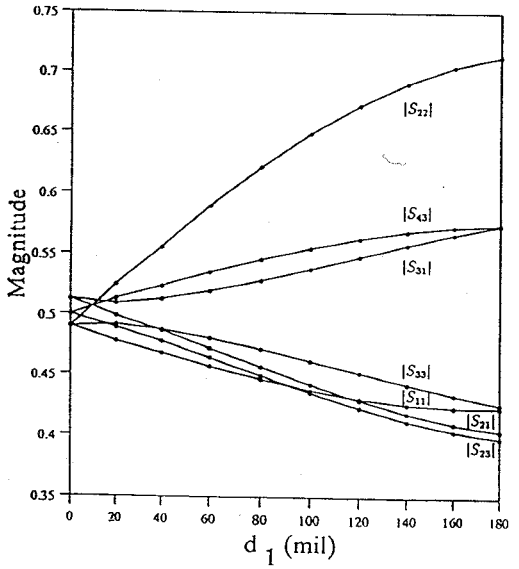
The learning procedure first uses the input vector to produce its own output vector and then compares this with the desired output. If there is no difference, no learning takes place. Otherwise the weights are changed to reduce the difference. It is noted that the desired output needs to be normalized within the range $[0 : 1]$ due to the use of the activation function in (2). The scheme for updating the weights is to achieve the minimization of the cost function defined as

$$E = \sum E_p = \frac{1}{2} \sum_p \sum_k (t_{pk} - o^{out}_{pk})^2, \quad (3)$$

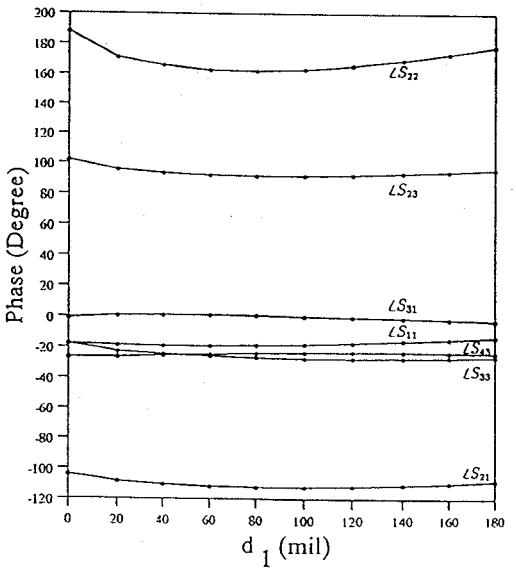
where t_{pk} and o^{out}_{pk} are the kth element of the desired output and observed output respectively by the presentation of the pth input sample. The amount of the weight change is proportional to the derivative of the cost function with respect



(a)



(b)



(c)

Figure 2: (a) The junction geometry. (b) Magnitude and (c) Phase of the S parameters

to each weight. Thus,

$$\Delta_p w_{ji}^a \propto -\frac{\partial E_p}{\partial w_{ji}^a} = i_{pi} f'(h_{pj}^{in}) \sum_k \delta_{pk} w_{kj}^b \tag{4}$$

and

$$\Delta_p w_{kj}^b \propto -\frac{\partial E_p}{\partial w_{kj}^b} = h_{pj}^{out} f'(o_{pk}^{in}) \delta_{pk}, \tag{5}$$

where

$$\delta_{pk} = (t_{pk} - o_{pk}^{out}) f'(o_{pk}^{in}). \tag{6}$$

Eqs. (4)-(6) give a recursive procedure for computing the weight changes in the network according to a back-propagation scheme [4].

V. Simulation Results

For any given sampling space, the learning procedure will eventually lead to a solution. However, the accuracy of the solution depends on what samples we choose and how many samples we have. Fig. 4 shows a comparison of using different size of sampling space in the corporate feed problem. The case presented here was never chosen as a sample to train any network such that we might use it to test how close the networks approach the mapping. In Fig. 4, the solid line represents the array pattern calculated from the output currents based on a full-wave numerical result, while the other three curves are simulation results using neural networks with the number of training samples equal to 432, 1024 and 2400 respectively. It can be seen that the simulation results using neural networks can approach a full-wave result better as the number of samples increases. Finally, we used the well trained neural network to find the distance parameters associated with the 20dB, 25dB and 30dB Dolph-Chebyshev array patterns. The values of parameters and the patterns are shown in Table 1 and Fig. 5 respectively.

REFERENCES

[1] S.A. Amari, "Neural theory of association and concept formulation," *Biological Cybernetics*, 26, 175-185.

[2] J.A. Freeman and D.M. Skapura, "Neural networks: algorithms, applications, and programming techniques," Addison Wesley, 1991.

[3] R. Hecht-Nielsen, "Neurocomputing," Addison Wesley, 1990.

[4] D.E. Rumelhart, J.L. McClelland and the PDP research group, "Parallel distributed processing," The MIT Press, 1986.

[5] D.H. Ackley, G.E. Hinton, and T.J. Sejnowski, "A learning algorithm for Boltzmann machines," *Cognitive Science*, Vol. 9, pp. 147-169.

[6] D.H. Ballard, G.E. Hinton and T.J. Sejnowski, "Parallel visual computation," *Nature*, Vol. 306, pp. 21-26.

[7] P.B. Katehi and N.G. Alexopoulos, "Frequency-dependent characteristics of microwave discontinuities in millimeter-wave integrated circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-33, pp. 1029-1035, Oct. 1985.

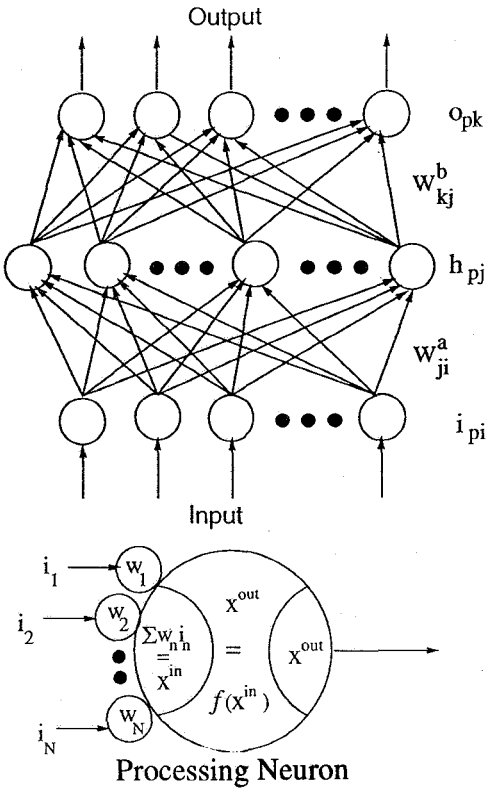


Figure 3: A three layer neural network.

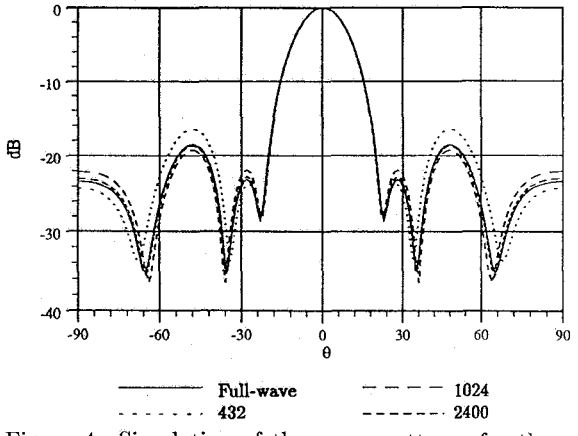


Figure 4: Simulation of the array patterns for the case of the corporate feed with the following distance parameters: d_1=60 mil, d_2=40 mil, d_3=120 mil, d_4=60 mil. d_5=120 mil.

d_1	d_2	d_3	d_4	d_5
20 dB Dolph-Chebyshev Array Pattern				
0mil	0mil	110mil	46mil	114mil
25 dB Dolph-Chebyshev Array Pattern				
38mil	40mil	0mil	76mil	124mil
30 dB Dolph-Chebyshev Array Pattern				
142mil	80mil	162mil	130mil	120mil

Table 1: Simulation results of the distance parameters for the 20dB, 25dB and 30dB Dolph-Chebyshev array sum patterns.

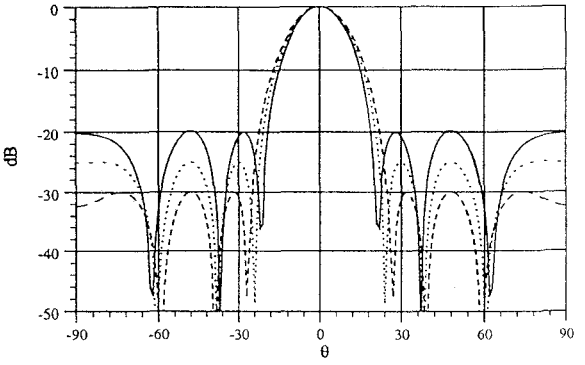


Figure 5: Simulation of the 20dB, 25dB, and 30dB Dolph-Chebysehv array sum patterns.

[8] W.P. Harokopus and P.B. Katehi, "Characterization of microstrip discontinuities on multilayer dielectric substrates including radiation losses," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-37, pp. 2058-2066, Dec. 1989.

[9] R.H. Jansen, "The spectral-domain approach for microwave integrated circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-33, pp. 1043-1056, Oct. 1985.

[10] S.C. Wu, H.Y. Yang, N.G. Alexopoulos and I. Wolff, "A rigorous dispersive characterization of microstrip cross and T junctions," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-38, pp. 1837-1990, Dec. 1990.

[11] T.S. Horng, N.G. Alexopoulos, S.C. Wu and H.Y. Yang, "Full-wave spectral-domain analysis for open microstrip discontinuities of arbitrary shape including radiation and surface-wave losses," *Int. J. Microwave Millimeter-Wave Computer-Aided Eng.*, Vol. 2, No. 4, pp. 224-240, 1992.

[12] T.S. Horng, W.E. McKinzie and N.G. Alexopoulos, Full-wave spectral-domain analysis of compensation of microstrip discontinuities using triangular subdomain functions," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-40, pp. 2137-2147, Dec. 1992.

[13] J.R. Mosig, "Arbitrarily shaped microstrip structures and their analysis with a mixed potential integral equation," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-36, pp. 314-323, Feb. 1988.

[14] K.A. Michalski and D. Zheng, "Analysis of microstrip resonators of arbitrary shape," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-40, pp. 112-119, Jan. 1992.

[15] X. Zhang and K.K. Mei, "Time-domain finite difference approach to the calculation of the frequency-dependent characteristics of microstrip discontinuities," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-36, pp. 1775-1787, Dec. 1988.

[16] D.M. Sheen, S.M. Ali, M.D. Abouzahra and J.A. Kong, "Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip discontinuities," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-38, pp. 849-857, July 1990.