Progress In Electromagnetics Research, PIER 69, 67–75, 2007

LTCC INTERCONNECT MODELING BY SUPPORT VECTOR REGRESSION

L. Xia, R. Xu, and B. Yan

College of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu, China

Abstract—In this paper, we introduce a new method: support vector regression (SVR) method to modeling low temperature co-fired ceramic (LTCC) multilayer interconnect. SVR bases on structural risk minimization (SRM) principle, which leads to good generalization ability. A LTCC based stripline-to-stripline interconnect used as example to verify the proposed method. Experiment results show that the developed SVR model perform a good predictive ability in analyzing the electrical performance.

1. INTRODUCTION

LTCC technology is growing rapidly as one of the suitable 3D packaging technology for size and cost reduction. Accurate modeling of 3D complex structures (e.g, LTCC multilayer interconnect structure) and fast and effective design tools are needed for the design of microwave circuits. Full-wave electromagnetic (EM) simulations are, typically computational accuracy but time-consuming, especially for the design adjustment and optimization. The methods of modeling based on sample data such as artificial neural network (ANN) are popular applied in recent years for its non-linear functional approximation property [1–3]. However, ANN modeling method depends on the network's structure and the complexity of the samples, which may cause over-fitting and low generalization ability [4].

Vapnik's support vector machine (SVM) theory [5] has been successfully applied for classification and regression problems [6, 7]. Many research results show that SVM has crucial advantages. Firstly, SVM solves a convex constrained quadratic optimization problem, whose error surface is free of local minima and has a unique global optimum. Secondly, SVM approach based on structural risk minimization (SRM) principle instead of empirical risk minimization (ERM) which used in ANN approach. SRM principle implements well trade-off between the model's complexity and its generalization ability [8]. Furthermore, support vector machine is based on small-sample statistical learning theory, whose optimum solution is based on limited samples instead of infinite sample. Commonly, support vector machine regression tasks called as support vector regression (SVR). In this paper, the usefulness of the introduced method is verified using an example.

2. SUPPORT VECTOR REGRESSION

Given a training dataset (y_i, x_i), $i = 1, 2, \ldots, n$. n is the size of training data. SVR tries to find the mapping function $f(x)$ between the input variable and the desired output variable. In formula this read as:

$$\begin{aligned} f(x) &= \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) K(x_i, x) + b \\ &= \sum_{i=1}^{n} (\alpha_i^* - \alpha_i) \phi(x_i) \phi(x) + b \\ &= w\phi(x) + b \end{aligned} \tag{1}$$

$\alpha_i^* \geq 0$, $\alpha_i \geq 0$, $(i = 1, 2, \ldots, l)$ are the Lagrange multipliers. $\phi(x)$ is a vector representing the actual no-linear mapping function. $K(x_i, x) \equiv \phi(x_i)\phi(x)$ is the kernel function. $w = \sum_{i=1}^{n} (\alpha_i^* - \alpha_i)\phi(x)$, b is the offset of the regression function. Traditional regression method find the regression function $f(x)$ by the rule of empirical risk minimization principle, i.e.,
minimize:

$$R_{emp}[f] = \frac{1}{n} \sum_{i=1}^{n} L(f(x_i) - y_i) \tag{2}$$

with $L(f(x_i) - y_i) = |y - f(x)|_\varepsilon = \max|0, |y - f(x)| - \varepsilon|$.

$L(f(x_i) - y_i)$ represents the error function, ε is the insensitive loss function. y_i is real value, $f(x_i)$ is the prediction value. However, the actual risk minimization can not be realized only with the empirical risk minimization [5]. A typical example is the over-fitting of NN. Support vector regression method based on structural risk minimization principle (SRM), which minimize the following cost function:

$$\frac{1}{2}\|w\|^2 + C \cdot R_{emp}[f], \tag{3}$$

where $\frac{1}{2}\|w\|^2$ is the term characterizing the modeling complexity. C is a regularization which determines the trade off between model complexity and empirical loss function. After some reformulations and introduction of the slack variables: ξ_i, ξ_i^*. Equation (2) is transformed into primal problem:
minimize:

$$\frac{1}{2}\|w\|^2 + C \cdot \frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{4}$$

subject to:

$$\begin{cases} (w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i \\ y_i - (w \cdot x_i) - b \leq \xi + \xi_i^* \\ \xi_i > 0,\ \xi_i^* > 0,\ \varepsilon > 0. \end{cases}$$

According to [9], an improved SVR has been presented, Equation (4) can be changes to minimize:

$$\frac{1}{2}\|w\|^2 + C\left(\mu\varepsilon + \frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*)\right) \tag{5}$$

μ is added as a constant along with ε to modulate the model complexity and slack variables.

According to Wolfe dual theory, the Equation (5) can be reformulated into the primal problem:

$$W(\alpha, \alpha^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, x_j) \tag{6}$$

subject to :

$$\begin{cases} \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \\ \alpha_i \in [0, \frac{C}{n}] \\ \alpha_i^* \in [0, \frac{C}{n}] \\ \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) \leq C \cdot \mu \end{cases}$$

α_i, α_i^* can be obtained by maximization the Equation (6). According to Equation (1), parameter b can be obtained by using training data. Then the output can be predicted by Equation (1) for every new input parameter value.

3. EXPERIMENT AND RESULTS

In this section, a LTCC-based stripline to stripline interconnect used as an example to verify the method proposed in this paper. As shown in Fig. 1, LTCC substrate fabricated using twelve-layer Ferro A6 tape systems, which has a relative dielectric constant of $\varepsilon_r = 5.7$, $\tan\delta = 0.002$. Each fired single layer has the thickness of 0.1 mm. The stripline has the line-width of 0.2 mm and height of 0.6 mm to keep 50 Ω characteristic impedance. The variable design parameters are the radius of the metal via hole (r_{via}), open hole in the middle ground (r_{gnd}) and the radius of the metal pad (r_{pad}). All other physical dimensions are fixed. The operation frequency is also used as the input parameter. The port reference plane and physical layout parameters of the vertical interconnect are shown in Fig. 1.

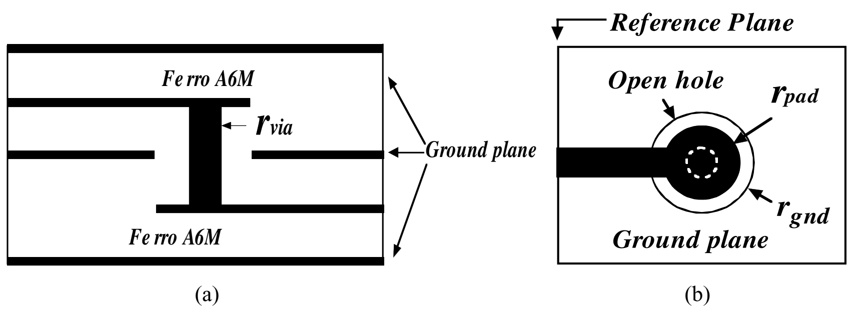

Figure 1. Structure of interconnect. (a) Cross section (b) Top view.

Similar to the ANN model, SVR estimates the non-liner function that encodes the fundamental interrelation between a given input and its corresponding output in the training data that acquired from EM simulation. This developed model then can be used to predict outputs for given inputs that were not included in the training data. We introduce two parameters a and b, let

$$r_{gnd} = a + r_{via}, \tag{7}$$

$$r_{pad} = b + r_{via}. \tag{8}$$

The input vector is

$$X = (a, b, r_{via}, f) \tag{9}$$

The output vector is

$$Y = (|S11|, |S21|). \tag{10}$$

The range of input parameters is given in Table 1, while the training and testing data are obtained by simulating in HFSS with these input variables separately. In this example, we only interested in the amplitude of the S parameters. Therefore, the amplitude of return loss ($|S11|$) and the insertion loss ($|S21|$) are computed as the model output parameters.

Table 1. Range selection of SVR variables.

Data / Variable	Training Data			Testing Data		
	Min	Max	Step	Min	Max	Step
r_{via} (mm)	0.05	0.15	0.05	0.075	0.125	0.05
a (mm)	0.05	0.25	0.1	0.1	0.2	0.1
b (mm)	0.05	0.25	0.1	0.1	0.2	0.1
f (GHz)	1	16	1	1	16	0.5

SVR calculation is performed using the SVM toolbox developed by C. C. Chang and C. J. Lin [10]. All programs are implemented in Matlab V6.5 (The Mathworks, Inc.) and carried out on an Intel Pentium IV 2.8 GHz with 1 GB of memory and running Windows XP.

Before running LIBSVM code, some SVR parameters need to be determined: the constant defining of kernel function (γ), tolerance of termination criterion (ε), the penalty parameter (C) and the constant ν. $v \in [0, 1]$ is the parameter to control the number of support vectors. After performing several experimentations with different variable value, the variables are fixed as: $\varepsilon = 0.00001$, $v = 0.1$, $C = 500$, and γ with the default value: $\gamma = 1/k$, k means the number of EM-SVR model input parameters. Moreover, the quality of model is evaluated as its prediction accuracy with mean squared error (MSE) and the liner correlation coefficient (R) [11]:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2 \tag{11}$$

$$R = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{12}$$

where x_i and y_i denote the EM simulated s-parameters value and the SVR predicted value and N is the number of validation data respectively. $\bar{x}$ is the EM simulated s-parameters samples mean and $\bar{y}$ is the v-SVR predicted samples mean value.

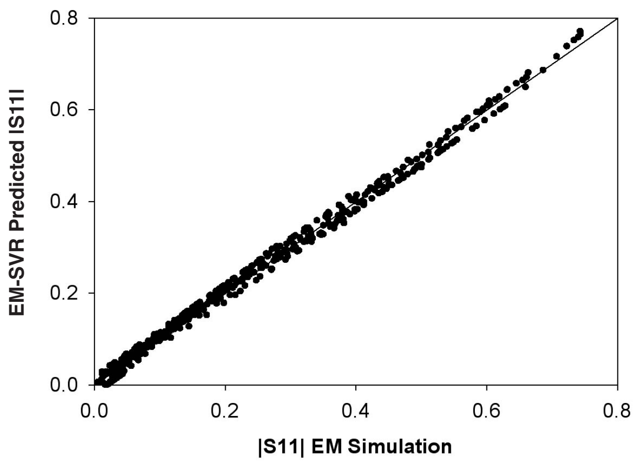

Figure 2. Scatter plots of EM simulated and SVR predicted $|S11|$ (Training data).

The scatter plots of the EM-SVR model predicted S-parameter compare to the training dataset and test dataset are shown in Fig. 2 ∼ Fig. 5. The y-axis is EM-SVR predicted value, while the abscissa is real value (EM simulated value). When EM-SVR predicted value equal to real value, the point lies in the diagonal line. That is to say, the more the points are concentrated around the diagonal line, the better the prediction precision are. Summary of MSE and correlation coefficient R are shown in Table 2. As can be seen from these results, excellent agreement between the EM-SVR model and electromagnetic simulation can be arrived.

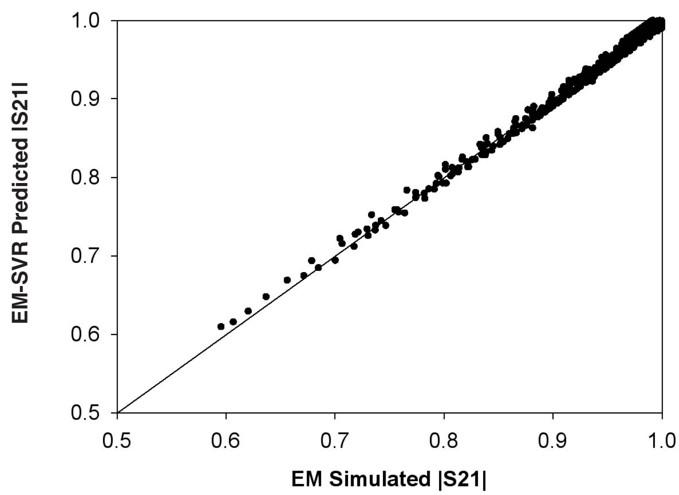

Figure 3. Scatter plots of EM simulated and SVR predicted $|S21|$ (Training data).

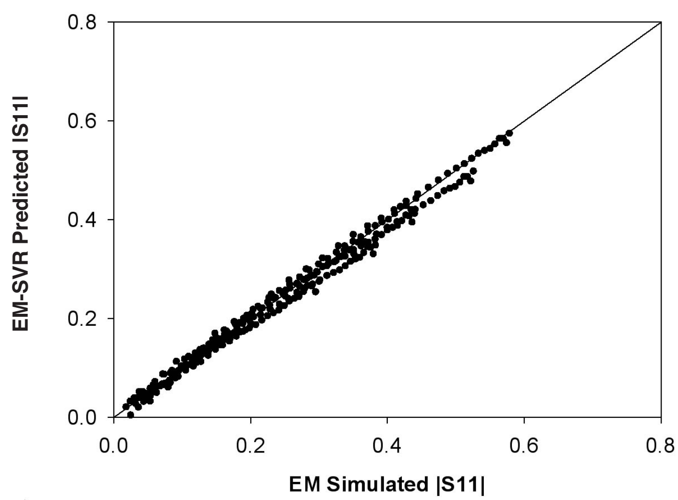

Figure 4. Scatter plots of EM simulated and SVR predicted $|S11|$ (Testing data).

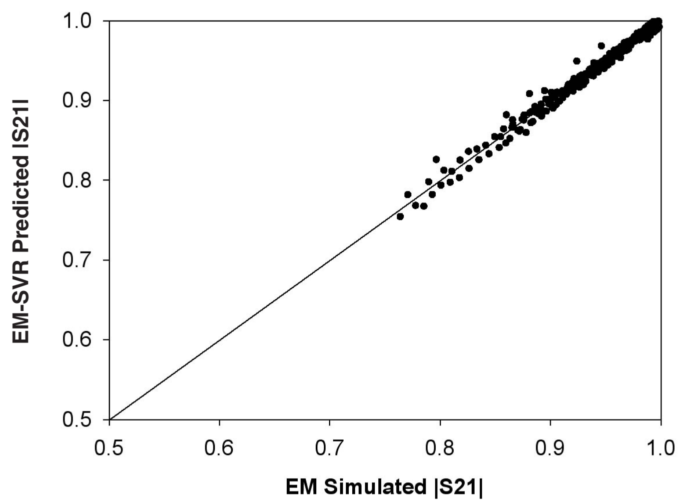

Figure 5. Scatter plots of EM simulated and SVR predicted $|S21|$ (Testing data).

Table 2. Correlation and error results for the EM-SVR model.

Data	Training Data		Testing Data	
$\|S\|$	$\|S11\|$	$\|S21\|$	$\|S11\|$	$\|S21\|$
MSE	$1.7695 \cdot 10^{-4}$	$4.1424 \cdot 10^{-5}$	$2.3153 \cdot 10^{-4}$	$6.3205 \cdot 10^{-5}$
R	0.9975	0.9968	0.9952	0.9908

4. CONCLUSION

Support vector machine regression method using in microwave transition modeling has been presented in this paper. To verify the method, modeling of a LTCC based stripline-to-stripline interconnect has been developed. The results verified the approach and the model by comparing with electromagnetic simulation results. The developed SVR model preserves the accuracy of the EM simulation, and it useful for interactive CAD of millimeter wave circuit design.

REFERENCES

1. Bermani, E., S. Caorsi, and M. Raffetto, "An inverse scattering approach based on a neural network technique for the detection of dielectric cylinders buried in a lossy half-space," *Progress In Electromagnetics Research*, PIER 26, 67–87, 2000.
2. Jin, L., C. L. Ruan, and L. Y. Chun, "Design E-plane bandpass filter based on EM-ANN model," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 8, 1061–1069, 2006.
3. Lee, Y. and D. S. Filipovic, "ANN based electromagnetic models for the design of RF MEMS switches," *IEEE Microwave and Wireless Components Letters*, Vol. 15, 823–825, 2005.
4. Chen, K., C. Ho, and H. Shiau, "Application of support vector regression in forecasting international tourism demand," *Tourism Management Research*, Vol. 4, 81–97, 2004.
5. Vapnik, V., *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
6. Wei, C., J. O. Chong, and S. S. Keerthi, "An improved conjugate gradient scheme to the solution of least squares SVM," *IEEE Trans. Neural Network*, Vol. 6, 498–501, 2005.
7. Shevade, S. K., S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to the SMO algorithm for SVM regression," *IEEE Trans. Neural Network*, Vol. 11, 1188–1193, 2000.
8. Bermani, E., A. Boni, A. Kerhet, and A. Massa, "Kernels evaluation of SVM based estimatiors for inverse scattering problems," *Progress In Electromagnetics Research*, PIER 53, 167–188, 2005.
9. Scholkopf, B., A. J. Smola, R. Williamson, and P. Bartlett, "New support vector algorithms," NeuroCoLT2 Technical Reports Series: NC2-TR-1998-031, 1998.
10. Chang, C. C. and C. J. Lin, "LIBSVM: a library for support vector machines," System documention, National Taiwan University, 2004.
11. Scheaffer, R. and J. Mcclave, *Statistics for Engineers*, Duxbury Press, Boston, 1982.