Journal of the European Ceramic Society 22 (2002) 1867–1872

www.elsevier.com/locate/jeurceramsoc

Investigation of $BaTiO_3$ formulation: an artificial neural network (ANN) method

Dong Guo[a,*], Yongli Wang[a], Juntao Xia[b], Cewen Nan[a], Longtu Li[a]

[a]*Department of Materials Science and Engineering, Tsinghua University, 100084 Beijing, PR China*
[b]*School of Chemical Engineering and Materials, Beijing Institute of Technology, 100081 Beijing, PR China*

Received 1 August 2001; accepted 11 November 2001

Abstract

Artificial neural networks (ANNs) are relatively new computational tools and their inherent ability to learn and recognize highly non-linear and complex relationships makes them ideally suited in solving a wide range of complex real-world problems. However, very few is known of the use of this technique in ceramics although it is often invoked in diverse areas in chemistry. Here application of ANN technique to model the $BaTiO_3$ based dielectric ceramic formulation was carried through. Based on the homogenous experimental design the experimental results of 21 samples were analyzed by a three-layer back propagation (BP) network. Through comparison we found that the ANN model is much more accurate than conventional multiple nonlinear regression analysis (MNLR) model for the same set of data. The results of ANN model were also expressed and analyzed by intuitive graphics. It indicates that the three-layer BP network based modeling is a very useful tool in dealing with problems with serious non-linearity encountered in the formulation design of dielectric ceramics.

Keywords: Neural networks; $BaTiO_3$; Algorithm; Capacitor; Dielectric properties

1. Introduction

Due to its temperature-stable property and harmlessness to the environment both in production and in application, $BaTiO_3$-based dielectric is now under fast development and is commonly used as multilayer ceramic capacitors (MLCC) of EIA X7R specification.[1] In order to satisfy different requirements and to enhance the dielectric performance of this kind of material, addition of different dopants is usually inevitable. Thus, the materials often have complex ingredients and it is usually difficult to explain the functions of the additives in the system. The ever-increasing need to discover ceramic formulation with good dielectric properties requires considerable precise mathematical models in general. It would be intractable for us to develop a reliable modeling for the ceramic compounding without enough knowledge before doing the job by using a conventional method.[2] Computer modeling is becoming more and more important to scientists and engineers in allowing them a way to "picture" what might be happening. ANNs are relatively new computational tools that have found extensive utilization in solving many complex real-world problems. The attractiveness of ANNs comes from the remarkable information processing characteristics of the biological system such as nonlinearity, high parallelism, robustness, fault and failure tolerance, learning, ability to handle imprecise and fuzzy information, and their capability to generalize.[3] Artificial models possessing such characteristics are desirable because (i) nonlinearity allows better fit to the data, (ii) noise-insensitivity provides accurate prediction in the presence of uncertain data and measurement errors, (iii) high parallelism implies fast processing and hardware failure-tolerance, (iv) learning and adap-tivity allow the system to update (modify) its internal structure in response to changing environment, and (v) generalization enables application of the model to unlearned data.[4] In the past decade, there have been numerous successful applications of ANN technique in diversified areas of science and engineering including pattern recognition and classification, voice and image processing, prediction,[5] digital communications,[6] and nonlinear system identification and control.[7] However, very little is known of the use of this technique in the investigation of ceramics.

In this study ANN technique is used to model the dielectric properties of $BaTiO_3$ based dielectric ceramics.

* Corresponding author. Fax: +86-10-6277-1160.
E-mail address: daguadong@sina.com (D. Guo).

0955-2219/02/$ - see front matter
PII: S0955-2219(01)00501-5

In order to compare the accuracy of this method with a conventional method, the same set of experimental data were also used to develop the MNLR model. Further effort to extend this method to other functional ceramics is continuing in the lab.

2. Fundamentals of ANN

ANNs are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience, not from programming. An ANN is formed from hundreds of single units, artificial neurons or processing elements (PE), connected with coefficients (weights), which constitute the neural structure and are organized in layers. The power of neural computations comes from connecting neurons in a network. Each PE has weighted inputs, transfer function and one output. The behavior of a neural network is determined by the transfer functions of its neurons, by the learning rule, and by the architecture itself. The weighed sum of the inputs constitutes the activation of the neuron. The activation signal is passed through a transfer function to produce a single output of the neuron. The transfer function introduces non-linearity to the network. During training, the inter-unit connections are optimized until the error in predictions is minimized and the network reaches the specified level of accuracy. Once the network is trained and tested it can be given new input information to predict the output. Although ANNs are drastic abstractions of the biological counterparts, the idea of ANNs is not to replicate the operation of the biological systems but to make use of what is known about the functionality of the biological networks for solving complex problems.[4]

Nielson[8] had proved in 1989 that a continuous function in a closed interval must be approximated using a three-layer feed forward network with back propagation (BP) of errors. The ANN employed here is a BP network consisting of three layers, i.e. input layer, hidden layer and output layer. The input and output layers consist of some neurons where the input information (e.g. component%) is presented to the network and some neurons where the response of the network (e.g. permittivity calculated) is registered, respectively. There is a hidden layer consisting of individual processing units between them. Both input and hidden layers have an additional node named bias neuron.

Processing proceeds from the input neurons to the output neurons via the hidden neurons. Except for the bias neuron, each of them in the network is linked with those neurons in the adjacent layer. The bias neuron connects only with those neurons in the upper layer. The degree of influence is dictated by connection weight being adjusted during training. Key steps for calculations performed in the network can be summarized as follows:

(1) Input some data x_i to the neurons in the input layer.

(2) Calculate the outputs from the hidden layer by the transfer Eqs. (1) and (2):

$$y_i = \sum w_{ij} x_i + \theta \tag{1}$$

$$H_j = 1/[1 + \exp(-\alpha y_j)] \tag{2}$$

where w_{ij} is the connection weight between the neurons i and j, θ is the bias or threshold value for neuron j that can be regarded as the nonzero offset in the data, H_j is the output of neuron j, and α is a parameter which expresses the non-linearity of the neuron's operation.

(3) Calculate the output O_k (the parameters to be studied) at the output neuron k by equations similar to Equs. (1) and (2).

(4) Calculate the correction factor (error) for all weights in the output layer using its output value O_k and the target output t_k:

$$\delta_k(t_k - o_k)o_k(1 - o_k) \tag{3}$$

(5) Update weights on the output layer by Eqs. (4) and (5):

$$w_{jk}^{\text{new}} = w_{jk}^{\text{old}} + \Delta w_{jk}(p) \tag{4}$$

$$\Delta w_{jk}(p) = \eta \delta_k H_j + \mu \Delta w_{jk}(p-1) \tag{5}$$

where Δw_{ij} is the correction of the weight between hidden layer neuron j and output neuron k, p and p-1 refer to the present and previous cycles of correction, respectively. The empirical parameters, η, is called the learning rate and μ is called the momentum.

$$\delta_j = H_j(i - H_j) \sum \delta_k w_{jk} \tag{6}$$

(6) Calculate the correction factor δ_j for the hidden layer:
(7) Update weights w_{ij} on the hidden layer based on Eqs. (7) to (8):

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \Delta w_{ij}(p) \tag{7}$$

$$\Delta w_{ij}(p) = \eta \delta_j H_i + \mu \Delta w_{ij}(p-1) \tag{8}$$

(8) Return to the first step and repeat with a new input example.

The iteration continues until the overall error between calculated and target outputs is approaching to the preset error criteria.

The BP algorithm, as mentioned above, was realized through a software developed in Visual Basic 6.0 on the platform of Microsoft Windows 98.[9] The MNLR and the ANN computation were carried out through this software.

3. Experimental procedure

3.1. Experimental design and materials preparation

By adding Nb_2O_5, La_2O_3, Sm_2O_3, Co_2O_3 and Li_2CO_3 into $BaTiO_3$, Qi Li et al have obtained a formulation satisfying X7R specification.[10] Among the additives Co_2O_3 and Li_2CO_3 seem to be very critical in determining the performance of the system. Therefore this system was selected and the influences of Co_2O_3 and Li_2CO_3 on the performance of the system were mainly investigated. For knowledge acquisition the homogenous experimental design offers schemes for scientific experimental design within the range of interest. Twenty one formulations gained from homogenous experimental design[11] act as the training data set. Each was composed of six ingredients and $BaTiO_3$ content is fixed as 100%.

$BaTiO_3$ and five additives were weighed, mixed by ball milling and then after drying the powder was pressed into disks. After being fired at 1280 °C for 4 h in atmosphere the disks were paved with an Ag-electrode. Temperature dependence of permittivity and loss properties of the samples were measured by a HP4192A impedance analyzer in the temperature range of −55–125 °C. Room temperature dielectric loss tgδ, room temperature permittivity ε_{25}, maximum TCC (temperature coefficient of capacitance) K ($K=|\varepsilon-\varepsilon_{25}|_{max}/\varepsilon_{25}$) in the range of −55–125 °C are selected as the object output. The ultimate results of the samples are list in Table 1.

3.2. Knowledge acquisition of the $BaTiO_3$ system

Because $BaTiO_3$ content is fixed, the ANN model is composed of five neurons in the input layer and four neurons in the output layer. An appropriate number of hidden neurons is an important factor determining the network's performance. From trials using different numbers of hidden neurons, 5–10, the minimum in root mean square errors were obtained for 8 hidden neurons, therefore, the ANN model has an architecture as shown in Fig. 1. The initial connection weights are set to be random between −0.3 and +0.3. Within the neural network the learning rate, the momentum factor and the convergence error are set to be 0.15, 0.075 and 0.01, respectively.

Table 1 gives the training data set. Because a value near 0 or 1 will lead to the "inaction" of the network, the input data (all in mg) and the output data were normalized to give values between around 0.05 and 0.95 prior to training through Eqs. (9) and (10).

$$X'_i = (X_i - \bar{X}_i)/\sigma$$

$$\text{where, } \bar{X}_i = \sum_{i=1}^{p} X_i, \sigma = \sqrt{\frac{(X_i\bar{X}_i)^2}{p}}, X_i \tag{9}$$

is the original input data, and p is the number of training samples.

Table 1
Experimental results of 21 formulations

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
tgδ	0.0152	0.0112	0.0201	0.0118	0.0106	0.0068	0.0129	0.0069	0.0104	0.0095	0.0124	0.0153	0.0207	0.0175	0.0171	0.0143	0.0148	0.0175	0.0266	0.0106	0.1260
ε_{25}	1999	2235	2200	2470	2044	2294	2123	2537	2475	3142	2200	3005	2338	3736	3155	3195	3036	2802	3207	5995	5997
K	238	155	171	99	114	120	103	115	48	48	102	51	57	55	43	55	42	44	46	73	78

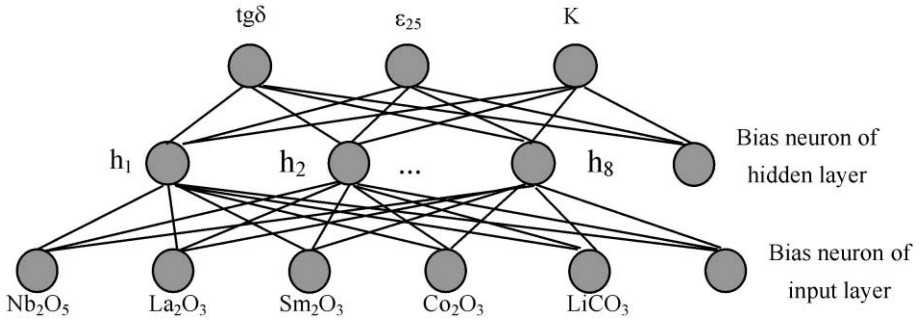

Fig. 1. Schematic view of the BP network for $BaTiO_3$ formulation model.

$$Y_i = 0.05 + 0.90 \times \frac{Y_i - Y_{\min}}{Y_{\max} - Y_{\min}} \quad (10)$$

where, Y_i is the orginal input data, and $Y_{\min}$, $Y_{\max}$ are the minimum value and maximum one in the orginal output data, respectively.

4. Results and discussion

4.1. Comparison of MNLR model and ANN model

In order to compare the accuracy between the ANN and conventional methods, The entire training data set (Table 1) was also used to develop MNLR models for correlation between the four parameters and $BaTiO_3$ formulation. For example, the ultimate model of $tg\delta$ and ε_{25} can be expressed in Eqs. (11) and (12), respectively.

$$\begin{aligned} tg\delta = {} & -0.2338 - (1.5799\mathrm{E}-03)^{*}Nb_2O_5 \\ & + (2.7261\mathrm{E}-03)^{*}La_2O_3 + 0.0332^{*}Sm_2O_3 \\ & - (3.3571\mathrm{E}-04)^{*}Co_2O_3 + (9.2345\mathrm{E}-03)^{*}Li_2CO_3 \\ & + (2.4387\mathrm{E}-04)^{*}Nb_2O_5^{*}La_2O_3 \\ & + (8.7886\mathrm{E}-05)^{*}Nb_2O_5^{*}Sm_2O_3 \\ & - (1.2898\mathrm{E}-03)^{*}La_2O_3^{*}Sm_2O_3 \\ & + (1.2593\mathrm{E}-04)^{*}Co_2O_3^{*}Li_2CO \end{aligned} \quad (11)$$

$$\begin{aligned} \varepsilon_{25} = {} & 7869.9063 - 234.0488^{*}Nb_2O_5 - 866.6985^{*}La_2O_3 \\ & - 266.8280^{*}Sm_2O_3 - 338.1283^{*}Co_2O_3 \\ & - 35.9705^{*}Li_2CO_3 + 78.8240^{*}Nb_2O_5^{*}La_2O_3 \\ & - 4.0792^{*}Nb_2O_5^{*}Sm_2O_3 + 57.5871^{*}La_2O_3^{*}Sm_2O_3 \\ & + 9.6048^{*}Co_2O_3^{*}Li_2CO_3 \end{aligned} \quad (12)$$

Where the additives are all expressed in mg and four reciprocal effects are considered based on our previous experiment results,[10] i.e. Nb_2O_5*La_2O_3, Nb_2O_5*Sm_2O_3, La_2O_3*Sm_2O_3, and Co_2O_3*Li_2CO_3. They contain 9 variables, five of which are independent. If other reciprocal effects are to be considered, the equations of the MNLR model will become more complicated. And the equations will become much more complicated while the accuracy of the model may show no change if the ingredients of the formulation increase.

Tables 2 and 3 present the connection weights between the input and hidden layers and the connection weights between the hidden and output layers, respectively. The calculated results comparing to observed results of ε_{25} of the 21 samples from MNLR and ANN models are illustrated in Figs. 2 and 3 respectively. Comparative predictions between MNLR and ANN models characterized through the root mean-square (RMS) error and the

Table 2
Connection weights between the input and hidden layers

	Nb_2O_5	La_2O_3	Sm_2O_3	Co_2O_3	Li_2CO_3	Input bias
h_1	−4.496	−1.468	−0.816	2.014	−2.23	6.426
h_2	1.317	0.492	0.495	0.168	−0.0662	1.398
h_3	4.829	−1.348	1.020	−0.510	−0.250	0.761
h_4	0.333	2.264	−3.141	0.723	−1.169	1.365
h_5	−2.076	2.987	−0.516	0.641	1.913	0.030
h_6	1.855	−2.959	2.705	−0.651	−1.647	−0.390
h_7	0.530	−0.800	−0.738	−0.474	−1.616	1.567
h_8	−2.038	−1.937	1.419	0.742	−2.524	−1.336

Table 3
Connection weights between the hidden and output layers

	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	Hidden bias
$tg\delta$	−4.306	0.444	0.544	−1.012	2.403	2.276	2.296	−2.342	−0.398
ε_{25}	−3.221	−6.664	1.081	2.472	−0.585	−1.184	−0.710	3.600	4.018
K	−0.500	−0.823	2.355	−2.229	2.084	0.093	−3.463	0.431	0.385

correlation coefficient are demonstrated in Table 4. Apparently, the ANN approach gives much better predictions than the traditional method.

4.2. Graphical analysis of the correlation between performance and formulation

By using a conventional method, in order to study the relationship between a component and a parameter of the formulation it is necessary to do some extra work. For example, if we want to study the influence of Co_2O_3 on the permittivity of the system we have to fix the concentration of other components and measure the parameters of the samples having different Co_2O_3 concentration. The original experimental data can not be utilized. However, by extracting the information of the ANN model registered, we can express the framework of the model in an intuitive way without any extra experiment. Here the graphical analysis capability of the ANN model is also illustrated. Based on our previous results, the ratio of Nb_2O_5: La_2O_3: Sm_2O_3 at about 3:2:1 may generate a good dielectric performance. Therefore we set the concentration of these three additives as 1.4 wt%, 0.8 wt% and 0.4 wt% respectively, then investigated the influence of Co_2O_3 and Li_2CO_3 on the properties of the system. Double components analysis of some parameters of the two additives were illustrated in Fig. 4–7. From Fig. 4 we can see that at different Co_2O_3 content ϵ_{25} shows a proportional relation with Li_2CO_3. Fig. 5 indicates that at different Li_2CO_3 content ϵ_{25} shows an inverse proportional relation with Co_2O_3. Fig. 6 and 7 suggest that increase of Li_2CO_3 will lead to the increase of TCC while increase of Co_2O_3 is helpful in depressing TCC. These results are consistent with our

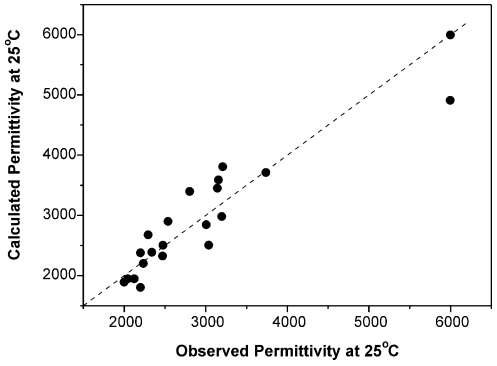

Fig. 2. Correlation between calculated and observed ε_{25} by MNLR model.

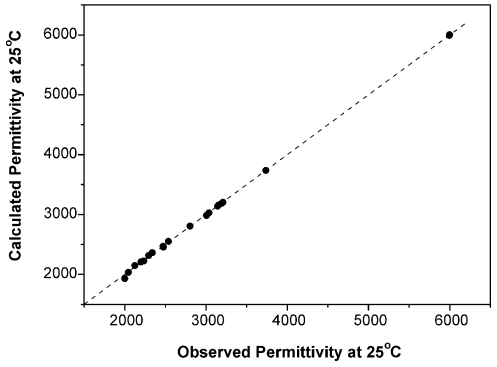

Fig. 3. Correlation between calculated and observed ε_{25} by ANN model.

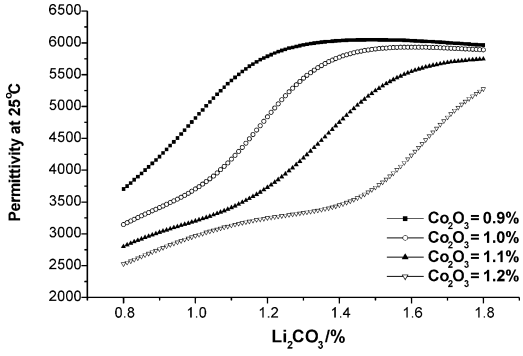

Fig. 4. Graphical analysis of the correlation between ε_{25} and Li_2CO_3.

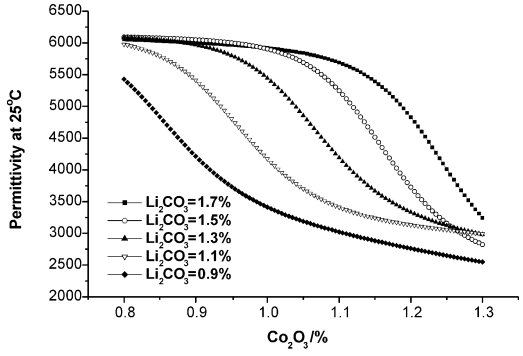

Fig. 5. Graphical analysis of the correlation between ε_{25} and Co_2O_3.

Table 4
Comparison of predictive ability between ANN and MNLR model

	MNLR model			ANN model		
	tgδ	ε_{25}	K	*tgδ*	ε_{25}	*K*
Correlation coefficient (R)	0.9835	0.9361	0.9086	0.9998	0.9998	0.9999
Root mean-square (RMS)	0.0044	382.78	20.93	0.0005	19.34	0.65

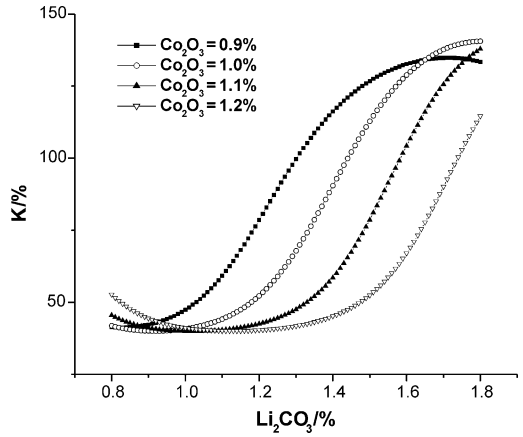

Fig. 6. Graphical analysis of the correlation between K and Li_2CO_3.

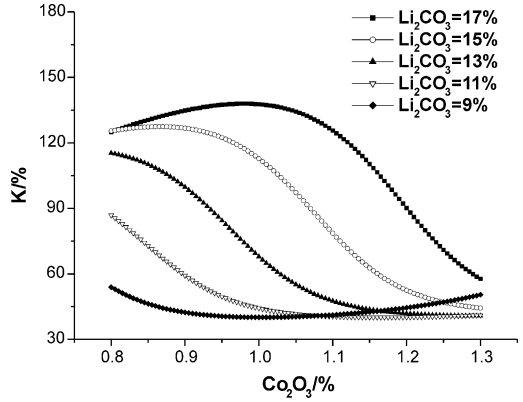

Fig. 7. Graphical analysis of the correlation between K and Li_2CO_3.

previous study.[10] It is difficult for us to find out an appropriate formulation having both high permittivity and low TCC. This is mainly caused by the addition levels of the additives selected. In fact, the addition level of Co_2O_3 and Li_2CO_3 are set to be ten times larger than those used in our previous study in order to precisely control the concentration of the two reagents in the samples.

5. Conclusion

A relatively new computational tools ANN, a methodology associated with nonlinear regression technique, is used here in investigating the dielectric performance of $BaTiO_3$ based dielectric ceramics. Application of a three-layer BP network modeling of the homogenous experimental results and comparison of the accuracy of the model with conventional MNLR model indicate that ANN is a very useful tool in dealing with problems with serious non-linearity encountered in complex formulation design of dielectric ceramics. It deserves more attention in the study of ceramics because of its unique ability to learn from and to adapt to their environment, and the ability to invoke weak assumptions about the underlying physical phenomenon responsible for the generation of the input data.

References

1. Kishi, H., Okino, Y. and Honda, M., The effect of MgO and rare-earth oxide on formation behavior of core-shell structure in $BaTiO_3$. *Jpn. J. Appl. Phys.*, 1997, **36**, 5954–5957.
2. Zupan J., Neural Networks for Chemists: An Introduction. New York, VCH, 1993.
3. Jain A. K., Mao J., Mohiuddin K. M., Artificial neural networks: a tutorial, Comput. IEEE, March 1996, 31-44.
4. Basheer, I. A. and Hajmeer, M., Artificial neural networks: fundamentals, computing, design and application. *J. Micro. meth.*, 2000, **43**, 3–31.
5. Haykin, S., *Neural Networks*. Maxwell Macmillan, Ontario, Canada, 1994.
6. Patra, J. C., Pal, R. N., Baliarsingh, R. and Panda, G., Nonlinear channel equalization for QAM signal constellation using artificial neural networks. *IEEE Trans. Syst. Man Cybernet. Part B: Cybernet*, 1999, **29**(2), 262–270.
7. Patra, J. C., Pal, R. N., Chatterji, B. N. and Panda, G., Identification of nonlinear dynamic systems using functional link artificial neural networks. *IEEE Trans. Syst. Man Cybernet. Part B: Cybernet*, 1999, **29**(2), 254–261.
8. Nielsen H., Theory of the back propagation neural network. In *Proc. of International Joint Conference on Neural Networks*, 1989, pp. 593–603.
9. Xia, Juntao, Huang, Congming and Wang, Jianqi, Development of expert system FRES 2.0 for flame retardant polymeric composites design. *Computers and Applied Chemistry (in chinese)*, 2000, **17**(5), 403–406.
10. Li Qi, Qi Jianquan, Wang Yongli, Gui Zhilun and Li Longtu, *J. Eur. Ceram. Soc.*, in press.
11. Fang, Kaitai, *Homogeneous Design and Homogeneous Design Tables*. Beijing, Science Press, 1994.