

# An Image Compression Scheme Based on Fuzzy Neural Network

Bo Wang\*, Yubin Gao

School of Computer and Control Engineering, North University of China, Taiyuan 030051, Shanxi, China

\*Corresponding author, e-mail: 22599870@qq.com

## Abstract

*Image compression technology is to compress the redundancy between the pixels to reduce the transmission broadband and storage space by using the correlation of the image pixels. Fuzzy neural network effectively integrates neural network technology and fuzzy technology; combines learning, self-adaptivity, imagination and identity and uses rule-based reasoning and fuzzy information processing in the nodes; thus greatly improving the transparency of fuzzy neural network. This paper mainly investigates the applications of fuzzy neural network in image compression and realizes the image compression and reconstruction of fuzzy neural network. It is demonstrated in the simulation experiment that the image compression algorithm based on fuzzy neural network has significant advantages in training speed, compression quality and robustness.*

**Keywords:** image compression, fuzzy theory, neural network

## 1. Introduction

Image compression technology refers to the method using as less as possible bits to represent the image signal from the signal source and possibly reducing such resource consumption occupied by the image data as the frequency bandwidth, storage space and transmission time etc., so as to transmit and store the image signals [1]. The main purpose of image compression is to eliminate the redundant information of images, including encoding redundancy, redundancy between pixels and psychological visual redundancy information [2]. In past decades, studies on the image compression have obtained the rapid development, and many effective algorithms have occurred, and compression standards such as JPEG, JPEG2000 and MPEG have been formed. In order to further compress images, we can start from two aspects: one is the use of vision system feature with the human eye as "final consumer" of the image information. The image compression based on human visual characteristics increasingly becomes the feature for people to study, and because of the complexity of the visual system, at present, there still exist a lot of unknown areas to be explored in this field. The second is the development of new compression tools and more intelligent algorithms. Owing to the excellent performance, there still exist a lot of room for the artificial neural network to be applied in the field of image compression [3].

Neural network almost is involved in every aspect of image processing, such as image segmentation, image enhancement, image pattern recognition, image restoration and image compression etc. Almost every kind of neural network can be directly or indirectly applied to the image compression, and the range of applications includes all sorts of damaged coding method or some key steps of these methods. What lot of neural network model realizes is a mathematical mapping from input space to the output space. The isomorphism between the image compression coding and neural network on the mathematical nature determines the fact that the neural network must have extensive application in the image compression coding field [4].

This paper firstly studies the theoretical basis and algorithm of the image compression, and then constructs the fuzzy neural network model and implementing steps of this algorithm, and finally under the Matlab environment, conducts the image compression and reconstruction through the experimental simulation. By comparative analysis, the algorithm in this paper has good convergence speed and high accuracy, effectively avoiding the defect existing in the image compression algorithm by the traditional neural network and improving the compression performance and the subjective quality of reconstructed image.

## 2. Basic Theory of Image Compression

### 2.1. Image compression mechanism

Because there are usually a large amount of data redundancy in the digital image, so it's possible to compress it to reduce the data representing the image, thus being convenient to store and transmit the image. Digital image is a two-dimensional functionsampled and quantified, and usually a two-dimensional real matrix is used to represent an image.

Take the image  $f(x, y)$  sampling for example, sample  $M, N$  times in the horizontal and vertical direction respectively, arrange these data into one matrix according to the relative position of sampling point and also quantify each element so as to get a numerical matrix. This matrix can be used to replace the function  $f(x, y)$ , that is to say, the digital image can be expressed by a matrix. Matrix element is called digital image pixel or pixel element. Representation form is as shown in formula 1:

$$f(x, y) \xrightarrow{\text{Sampling}} \begin{bmatrix} f(x_0, y_0) & f(x_0, y_1) & \cdots & f(x_0, y_{N-1}) \\ f(x_1, y_0) & f(x_1, y_1) & \cdots & f(x_1, y_{N-1}) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_{M-1}, y_0) & f(x_{M-1}, y_1) & \cdots & f(x_{M-1}, y_{N-1}) \end{bmatrix} \quad (1)$$

$$\rightarrow [f(i, j)]_{M \times N} \xrightarrow{\text{Quantification}} [f_i(i, j)]_{M \times N}$$

In which,  $f_i(i, j)$  represents the pixel value quantified.

If the sampling points are  $M, N$  and the quantified level is  $Q = 2^n$ , digits required to store a digital image is:

$$B = M \times N \times n \quad (2)$$

Image compression is to handle the numerical matrix and represent this numerical matrix with fewer data. Different amount of data are shown in different ways in terms of images, because different representation will produce data redundancy with different degree, and the purpose of image compression is to reduce these redundancy as soon as possible. As usual, there exist three basic data redundancy for the digital image: pixel redundancy, mental vision redundancy and encoding redundancy. When one or more of these three types of redundancy is reduced or eliminated, the image data compression will be realized [5].

### 2.2. Image compression problem analysis

For images, if quick or real-time transmission and large storage are required, the image data should be compressed. Under the equal communication capacity, if the image data is transmitted after the compression, more image information will be transmitted and also the communication ability will be enhanced. Image compression research is to find ways of high compression ratio and ensure that the compressed image has appropriate signal-to-noise ratio, and original signal should be restored after the compressed transmission, and besides, during the compression, transmission and restoration process, it is required that the image distortion shall be small so as to classify and recognize images easily.

The reason why images can be compressed lies in that the data volume of the original image is much greater than effective information amount it offers. That is to say, the original image data file contains a large number of redundancy and irrelevancy information. If  $D$  is used to represent the data volume and  $du$  the redundancy amount, the effective information amount  $I$  provided by the image is:

$$I = D - du \quad (3)$$

Usually the original image data is clear, and the redundancy is changing which varies according to the image using purpose. Difference between the "redundancy information" and "irrelevancy information" can be simply understood in this way: redundancy information is the

recurring information and its deletion exerts no loss to the original information; while the deletion of the irrelevancy information exerts some influence on the original information and does not affect the understanding of the information content under limiting conditions. But usually there is no need to accurately distinguish these two concepts. The algorithm only deleting the redundancy information is called "lossless compression", and it can fully restore original file according to the compressed file, but the lossless compression algorithm is lower in the compression. The algorithm deleting irrelevancy information is called "lossy compression", and it can only approximately restore the original file based on the compressed file, and its compression is very higher. Lossy compression can be adopted in most images according to the "fidelity" rule, and the lossless compression is only applied in some special conditions. Therefore, strategies shall be properly selected in line with the goal to win the largest benefit [6].

If  $n_1$  is used to represent the former data amount of one image and  $n_2$  the compressed data amount, the compression ratio  $C_r$  is defined as:

$$C_r = n_1 / n_2 \quad (4)$$

The redundancy amount  $R_d$  can be expressed as:

$$R_d = 1 - 1 / C \quad (5)$$

From a technical point of view, there are two ways to compress images. One is to reduce the total data amount being transmitted or stored by decreasing redundancy generated by each related image pixel, such as using discrete cosine transform (DCT) to reduce the correlation among data, and then keep the main component, thus reducing the amount of data. The other is to determine the appropriate coding schemes according to the dynamic data range and its occurrence frequency of, determine, and to realize the compression by arranging properly coding bits occupied by different data to reduce the overall bits required [7].

### 3. Fuzzy Neural Network

#### 3.1. Introduction to fuzzy neural network

##### 3.1.1. Fuzzy neural network mechanism

Fuzzy neural network is the process to handle samples fuzzily, and the samples should be constantly transformed into regular forms in the handling process. The mapping relation between input and output amount is represented by subordinate function. The input indicator vector is  $x = [x_1, x_2, \dots, x_n]^T$ ,  $x_i$  stands for the fuzzy variable. Suppose  $T(x_i) = \{A_i^1, A_i^2, \dots, A_i^{m_i}\}$  and  $i = 1, 2, \dots, n$ , in which, the value set of number  $j$  of  $x_i$  is  $A_i^j (j = 1, 2, \dots, m_i)$ , and suppose one certain fuzzy set of the domain  $U_i$  is  $A_i^j$ , and  $\mu_{A_i^j}(x_i) (i = 1, 2, \dots, n, j = 1, 2, \dots, m_i)$  is the subordinate function of  $x_i$  being subordinate to  $A_i^j$ . The output variable is  $T(y) = \{B_1, B_2, \dots, B_{m_y}\}$ , in which,  $B_j = \{j = 1, 2, \dots, m_y\}$  is the value of number  $j$  of  $y$ , and define the fuzzy set of domain  $U_y$  is  $B_j$ , and  $\mu_{B_j}(y)$  is the subordinate function of  $y$  being subordinate to  $B_j$  [8].

##### 3.1.2. Fuzzy neuron

Like ordinary artificial neural network, fuzzy neural network is usually composed of a large amount of fuzzy or non-fuzzy neurons that connect together according to certain topological structure. Here are three basic types of fuzzy neurons.

###### (1) Fuzzified neuron

The fuzzy neuron is a kind of neuron which can quantify or standardize the observation or input values. In simple terms, the role of fuzzy neuron lies in transforming input values into the fuzzy values:

$$y = F(x) \quad (6)$$

The input value  $x$  can either be discrete or continuous, definite or fuzzy. The output value is the value of the subordinate function  $F(\cdot)$  of one certain fuzzy set.

#### (2) Defuzzification neuron

Defuzzification neuron is the information processing unit that can transmit the output result represented by the "distribution value" in the form of "fixed value". The input-output relationship represented by the defuzzification is:

$$y = \varphi(x_1, x_2, \dots, x_n) \quad (7)$$

In which  $\varphi$  is the defuzzification function. Common defuzzification methods include the maximum-taking method and the centroid-taking method. The maximum-taking method is to extract the point value of the "distribution value" function at the maximum point as the definite output value. The centroid-taking method is to extract the value of the "distribution value" function at the centroid point as the definite value[9].

#### (3) Fuzzy logic neuron

Fuzzy logic neuron is the most important and most commonly used fuzzy neural network, and its input-output relationship is:

$$\begin{aligned} u &= I(x, w) \\ y &= f(u - \theta) \end{aligned} \quad (8)$$

In which,  $x = (x_1, x_2, \dots, x_N) \in [0, 1]^N$  is the neuron input,  $\omega = (\omega_1, \omega_2, \dots, \omega_N) \in [0, 1]^N$  is the neuron connection weight, and  $u$  is the neuron inner state,  $y$  is the neuron output,  $\theta$  is the neuron valve value,  $f$  is the monotonically increasing output function,  $I$  is the fuzzy logic function or fuzzy integrated function and its specific form is determined by actual condition and needs. For example, all the following functions can be taken as specific expression form:

•Weighted sum function:  $u = \sum \omega_i x_i$ .

Integration by taking small first and then large:  $u = \vee(\omega_i \wedge x_i)$ .

•Integration by quadrature first and then taking large:  $u = \vee(\omega_i x_i)$ .

•Integration by quadrature first and then taking large:  $u = \sum (\omega \wedge_i x_i)$ .

## 4. Establishment Of The Image Compression Algorithm Based On Fuzzy Neural Network

### 4.1. Fuzzy neural network structure

Typical fuzzy neural network structure is as shown in Figure 1:

(1) The first layer is input layer: the input layer node in fuzzy neural network is the entrance of the fuzzy information, transmitting the input information to the next layer. Each node in such layer respectively represents the input information  $x_i (i = 1, 2, \dots, n)$ , therefore, the value of nodes in input layer is determined by the dimension of the input information, and this is  $N_1 = n$ .

(2) The second layer is fuzzification layer: the role of this layer plays in the entire network is to calculate the subordinate function  $\mu_i^j (i = 1, 2, \dots, n; j = 1, 2, \dots, m_i)$ , in which,  $n$  is the dimension of the input amount. Each component has its corresponding node, that is to say, the node number  $m_i$  corresponds with the number of fuzzy classification  $x_i$ . The commonly used

Gaussian subordinate function is represented as  $\mu_i^j = \exp(-\frac{x_i - c_{ij}}{\sigma_{ij}})^2$ , in which  $c_{ij}$  and  $\sigma_{ij}$  are

the center and width of the subordinate function respectively. The total node number of this layer is  $N_2 = \sum_{i=1}^n m_i$ .

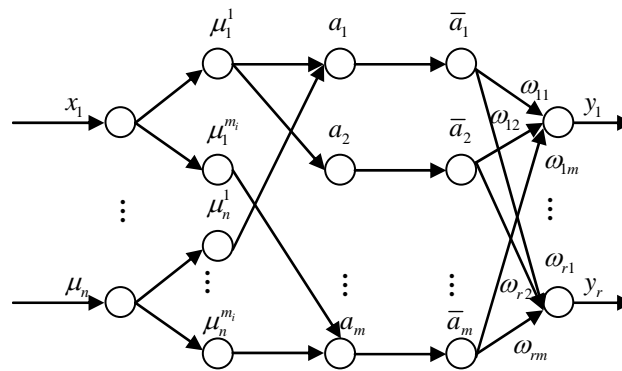


Figure 1. Fuzzy neural network structure

(3) The third layer is the fuzzy reasoning layer: each node in the fuzzy reasoning layer contains certain fuzzy rules. The data fuzzified match with the fuzzy rules of the fuzzy reasoning layer, and at the same time, the fitness of each fuzzy rule can be measured:

$$a_j = \mu_1^{i_1} \mu_2^{i_2} \wedge \mu_n^{i_n} \quad (9)$$

In which,  $i_n \in \{1, 2, \dots, m_n\}$ ,  $i_1 \in \{1, 2, \dots, m_1\}$ ,  $j = 1, 2, \dots, m$ ;  $m = \prod_{i=1}^n m_i$

The total nodes this layer is  $N_3 = m$ . The input variables are given, and the membership value only exists when the language variable value is nearly close to the input variable, and the membership value will be very small if the language variable value is quite far away from the input point. When membership degree is very small, such as less than 0.03, it can be approximated to 0, so, the output result  $a_j$  of most nodes is 0.

(4) The fourth layer is the normalized layer: node number of this layer is  $N_4 = N_3 = m$ . The function of this layer is to conduct normalized calculation towards the fitness value of the third layer:

$$\bar{a}_j = \frac{a_j}{\sum_{i=1}^m a_i}, j = 1, 2, \dots, m \quad (10)$$

(5) The fifth layer is the output layer: it is also called the anti-fuzzification layer. The clearness of the fuzzy neural network is realized in this layer.

$$y_i = \sum_{j=1}^m \omega_{ij} \bar{a}_j, i = 1, 2, \dots, r \quad (11)$$

In which,  $y_i$  is the result caused by the fuzzy neural network passes the output layer.

The learning parameter of fuzzy neural network includes two kinds: one is the  $c_{ij}$  value of subordinate function and its  $\sigma_{ij}$  value; the other is  $\omega_{ij}$  value of output layer of fuzzy neural network at the last time [10].

## 4.2. Establishment of the image compression algorithm based on fuzzy neural network

### 4.2.1. Basic idea

The basic idea using the mode conversion capability of multi-layer feed-forward network to realize the data transformation (code) is: map one group of input mode to one group of output mode through the middle layer (i.e. fuzzification layer, fuzzy reasoning layer and normalized layer), and make the output model equal to the input mode as soon as possible. The transformation of input layer and middle layer can be taken as the compressing and encoding process; and the transformation of the middle layer and output layer can be taken as the decoding process. Figure 2 gives a brief explanation of this idea.

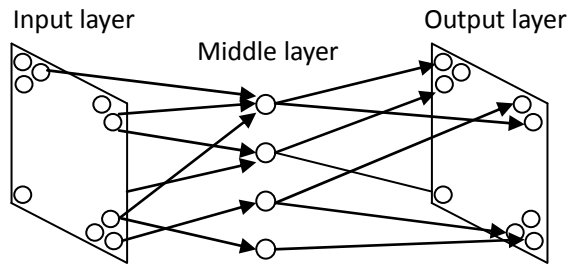


Figure 2. the basic idea of image compression based on fuzzy neural network

Assume that the network input layer and output layer are respectively composed of same  $M$  neurons, and neuron number  $K$  of the middle layer is smaller than  $M$ . Provide the same learning mode in the input layer and output layer (that is, the teacher mode is the input mode). After network studying, its underlying layer shall be able to give different encoding expression for each input modes among  $M$  input modes. Its basic idea is to make the original data pass the waist type network bottleneck, and expect to gain relatively compact data expression at the network bottleneck, in order to achieve the purpose of compression. In the process of network learning, adjust the network weights through training and make the reconstruction image similar with the training image possibly in mean error sense. The trained network can be used to perform the data compression task, and the weighted value between network input layer and the middle layer is equivalent to encoder. The original image data transmitted from the input end is processed by the fuzzy neural network to gain the output data in the middle layer, and the output data is the compression code of the original image, and the vector of the output layer is the reconstructed image data after the compression [11].

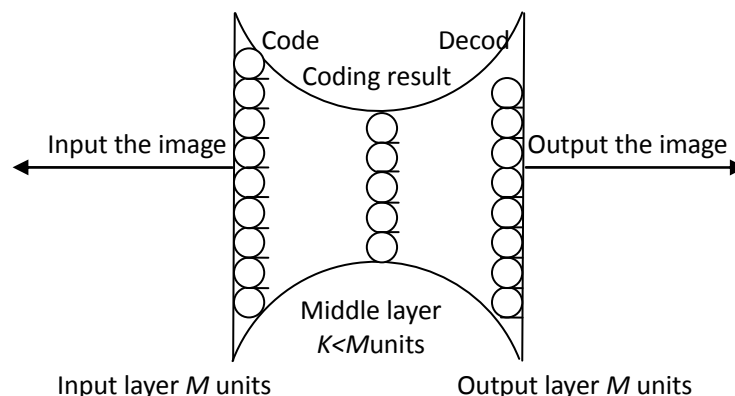


Figure 3. Schematic diagram of image compression based on fuzzy neural network

Network includes the input layer, middle layer and output layer. In the learning process, the image data is not only sent to the input layer, but also to the output layer as the teacher signals. When the network is properly trained, the process from the input layer to hidden layer is the network encoding process, and the process from hidden layer to output layer is the network decoding process. The continuous network training and network weight adjustment minimize make the network input and output mean square error, which eventually compresses the  $N$ -dimension vector into  $K$ -dimension vector ( $K < N$ ). The algorithm steps adopted the fuzzy neural network to compress image are as follows:

The algorithm flow chart is shown in Figure 4.

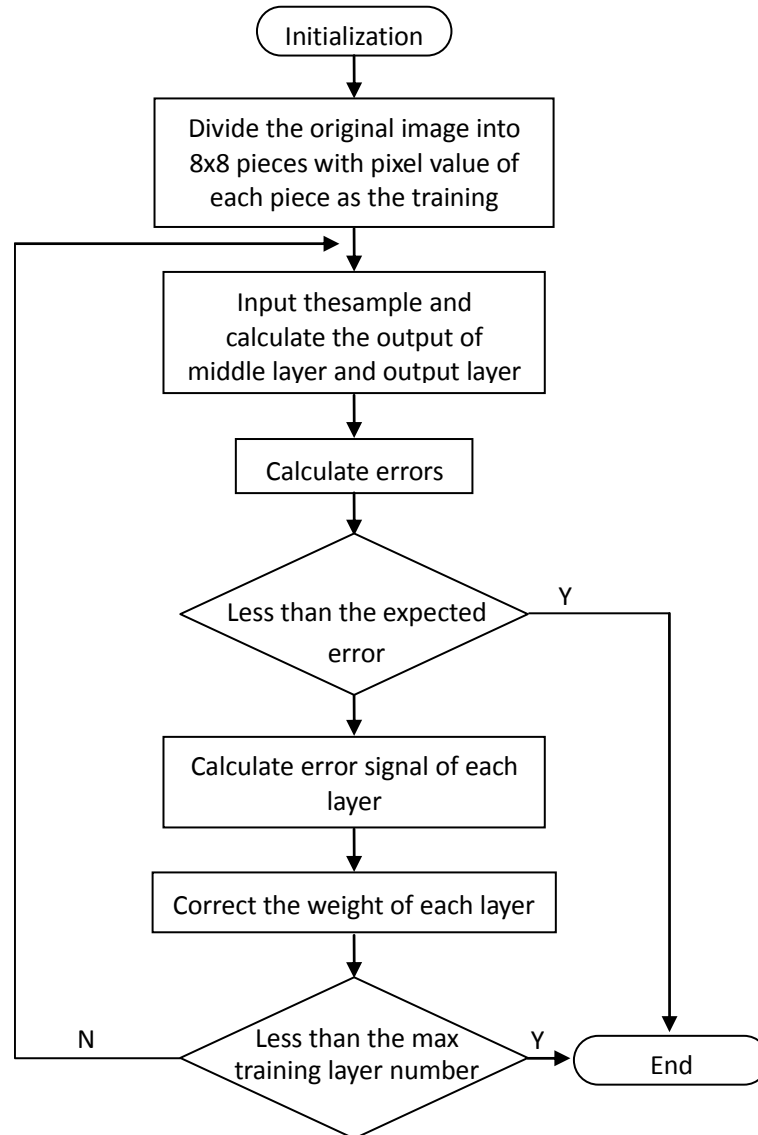


Figure 4. Flow chart of image compression based on fuzzy neural network

## 5. Experimental simulation and analysis

In order to demonstrate the effectiveness of this algorithm, we conduct a comparative contrast on the reconstructed image by Figure 5. The selected Lena image size is 512 x 512. Due to the image used is larger, 8x8 module is used in this experiment to improve the

compression efficiency. We can see that the compression method based on the fuzzy neural network is very clear and it is hard for the human eye to find any trace of distortion. While, the image based on traditional BP artificial neural network has obvious module effect and it looks that the whole image not smooth with the subjective feeling that the image is composed of "grids". This shows that the image compression based on the fuzzy neural network is stronger in the image compression capacity. The algorithm in this paper treats each coefficient as the function of the coordinate and also combines the learning, self-adaptiveness, imagination, recognition and fuzzy information processing. The fuzzy reasoning network training achieves the effect of function approximation, and the decoding end can decode directly only with the network weights, which effectively avoids the defect of traditional neural network in image compression algorithm and improves the compression performance and the subjective quality of reconstructed image.



(a) Original image



(b) Image compression method based on fuzzy neural network



(c) Image compression method based on BP neural network

Figure 5. Comparison of image compression based on different methods

## 6. Conclusion

Because most neural network models have strong ability to recognize and classify patterns, and this kind of pattern recognition and pattern classification ability provides a powerful tool to solve the pattern classification problem in image coding scheme. This paper studies deeply the theoretical basis, algorithm, network model and algorithm realization of the image compression based on the fuzzy neural network and also realizes the image compression and reconstruction to gain the higher compression ratio. In face of the current huge amounts of data storage, how to realize the situation that what we do is what we store and how to complete the demand analysis of human eye in the storage stage with view to largely compress the data, and



how to make limited space store more useful data, are what we should concern and try to accomplish.

### References

- [1] Roman S. New Simple and Efficient Color Space Transformations for Lossless Image Compression. *Journal of Visual Communication and Image Representation*. 2014; 25(5): 1056-1063.
- [2] Mario ARD, Hermilo SC. Refined Fixed Double Pass Binary Object Classification for Document Image Compression. *Digital Signal Processing*. 2014; 30(7): 114-130.
- [3] Jin-Yu Z, Wei Z, Zheng-Wei Y, Gan T. A Novel Algorithm for Fast Compression and Reconstruction of Infrared Thermographic Sequence Based on Image Segmentation. *Infrared Physics & Technology*. 2014; 67(11): 296-305.
- [4] Mahmood O. Fully Fuzzy Polynomial Regression with Fuzzy Neural Networks. *Neurocomputing*. 2014; 142(22): 486-493.
- [5] Kartik S, Ratan KB, Amitabha C. Image Compression Based on Block Truncation Coding using Clifford Algebra. *Procedia Technology*. 2013; 10: 699-706.
- [6] Jianji W, Nanning Z, Yuehu L, Gang Z. Parameter Analysis of Fractal Image Compression and Its Applications in Image Sharpening and Smoothing. *Signal Processing: Image Communication*. 2013; 28(6): 681-687.
- [7] Bo M, Yifang B. Generalization of 3D Building Texture using Image Compression and Multiple Representation Data Structure. *SPRS Journal of Photogrammetry and Remote Sensing*. 2013; 79(5): 68-79.
- [8] S Muralisankar, N Gopalakrishnan. Robust Stability Criteria for Takagi–Sugeno Fuzzy Cohen–Grossberg Neural Networks of Neural Type. *Neurocomputing*. 2014; 144(20): 516-525.
- [9] Choon KA. Receding Horizon Disturbance Attenuation for Takagi–Sugeno Fuzzy Switched Dynamic Neural Networks. *Information Sciences*. 2014; 280(10): 53-63.
- [10] Fayez FMES. Adaptive Hybrid Control System using A Recurrent RBFN-based Self-Evolving Fuzzy-Neural-Network for PMSM Servo Drives. *Applied Soft Computing*. 2014; 21(8): 509-532.
- [11] J Yang, H Shi, B Feng, L Zhao, C Ma, X Mei. Applying Neural Network Based on Fuzzy Cluster Pre-processing to Thermal Error Modeling for Coordinate Boring Machine. *Procedia CIRP*. 2014; 17: 698-703.