

## Wood Defects Recognition Based on Fuzzy BP Neural Network

Hongbo Mu<sup>1,3</sup>, Mingming Zhang<sup>1,2,3</sup>, Dawei Qi<sup>\*1</sup>, Shuyue Guan<sup>1</sup> and Haiming Ni

<sup>1</sup>College of Science, Northeast Forestry University, Hexing Road 26, Harbin,  
Heilongjiang Province, 150040, PR China  
[mhb-506@163.com](mailto:mhb-506@163.com)

<sup>2</sup> College of Bioinformatics Science and Technology, Harbin Medical University,

<sup>3</sup> These authors are the joint first authors

\*Corresponding Author, [qidw9806@126.com](mailto:qidw9806@126.com)

### Abstract

Firstly, we applied the X-ray non-destructive testing technology to detect wood defects for getting the images. After graying the images, we calculated their GLCMS(Gray Level Co-occurrence Matrixes), then we normalized GLCMS to obtain the joint probabilities of GLCMS. The feature vectors of images, which included 13 eigenvalues of images were calculated and extracted by the joint probability of GLCMS. The fuzzy BP neural network(abbreviated as FBP) was designed by combining fuzzy mathematics and BP neural network. And the FBP neural network was regarded as the membership function of feature vectors, the outputs of the network was regarded as the degree of membership to the feature vectors in each category. We use the maximum degree of membership method for the pattern recognition of feature vectors, so the automatic identification and classification for feature vectors were achieved, and then the automatic identification of wood defects was realized.

By simulated study and training many times, the results shown that the average recognition success rate of the network was more than 90%, and some FBP networks had an extremely high recognition success rate to training samples and test samples.

**Keywords:** wood defects; Gray Level Co-occurrence Matrixes (GLCM); feature extraction; fuzzy BP neural network; membership function

### 1. Introduction

Nowadays, forest resources reduce gradually, but the demand for wood is increasing, and the construction of the ecological environment needs human to reduce the wood cutting. Besides looking for substitute materials of wood, improving the efficiency of wood utilization also can effectively alleviate this contradiction. In the wood processing, wood defects have restricted the efficiency of wood utilization. To realize the recognition of wood defects and to classify wood defect categories are of great significance to improve the utilization rate of wood. Traditional artificial defects inspection depends on the experience and attention of testing inspectors. To the same defect, different inspectors with their experiences may get different conclusions. Additionally, the inspectors' fatigue can also affect the testing results. Hence, realizing automatic identification for wood defects not only can get rid of the dependence on people in the process of detection, but also can improve production efficiency. In recent years, although in the field of wood defect recognition researching work has made great progress, it not yet has reached the mature degree of industrialization. How to get the high accuracy intelligent detection is the key to solve the pattern recognition for wood defects, and also the wood processing industry needs to solve such the problem. In view of this, in this paper, three kinds of typical flaws: cracks, knots, and rotten defects are studied for automatic identification [1]. First of all, in the case of not damaging the

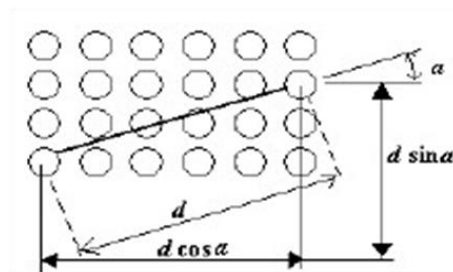
wood, we apply X-ray to obtain the external and internal texture images of the wood with defects [2,3]. Then, we use preprocessing and feature extraction to the wood defect images, combine the BP neural network with fuzzy mathematics, design the FBP fuzzy neural network [4,5], and realize the automatic recognition and classification of wood defects. These works lay a certain foundation for wood processing in the future, especially for the timbers with defects.

## 2. Gray Level Co-occurrence Matrixes

The generation of the Image Gray Level Co-occurrence Matrix, as showing in Fig.1, In the Gray image (Grayscale:  $N_g$ , Pixels:  $N_x \times N_y$ ), any point  $(x, y)$  and another point  $(x + d \sin \alpha, y + d \cos \alpha)$ , which the distance of deviation from it is  $d$ . And set the gray values of the two points for  $(i, j)$  -  $i, j \in \{1, 2, \dots, N_g\}$ . Let the point  $(x, y)$  move on the whole picture, we can get all kinds of the value of  $(i, j)$ , and the gray level co-occurrence matrix (GLCM) is obtained. In the matrix GLCM, the  $i$ -th row  $j$ -th column element  $G(i, j)$  is the number of occurrences for grey value of  $(i, j)$ . If we also include the gray series for  $N_g$  that does not appear in the image grey value into the gray level co-occurrence matrix, the matrix GLCM will be the  $N_g \times N_g$  phalanx. Although, by the gray level co-occurrence matrix we can analyze the image alignment rules and local mode characteristics *et al*, it generally does not directly apply for symbiotic matrix GLCM, but do secondary statistics based on it to get the secondary statistic values[6-8].

$$\text{Let } P(i, j) = \frac{G(i, j)}{R} \quad (1)$$

$R$  is the total number of all the pairs of pixels appeared, it is the normalized constant; in matrix  $P$ ,  $P(i, j)$  means the occurrence probability for grey value of  $i$ -th row  $j$ -th column element  $(i, j)$ . Namely,  $P$  is normalized probability matrix, and use the joint probability matrix  $P$  instead of the original gray level co-occurrence matrix (GLCM) for the feature extraction. In this way, the probability of two pixels gray scales occurred at the same time changes the space of the  $(x, y)$  coordinates into a gray level description of the  $(i, j)$ , and forms a joint probability gray level co-occurrence matrix[9,10].



**Figure 1. The Sketch of a Pair of Pixels**

Moreover, matrix  $P$  also affected by distance  $d$  and angle  $\alpha$ , is the function for  $d$  and  $\alpha$ . Different distances and angles will lead to get the different joint probability of gray level co-occurrence matrixes.

When  $\alpha = 0$ , a pair of pixels is horizontal, for a 0 - degree scan; When  $\alpha = 90$  degrees, a pair of pixels is vertical, for a 90 - degree scan; When  $\alpha = 45$  degrees, a pair

of pixels is right diagonal, for a 45 - degree scan; at the time of  $\alpha = 135$  degrees, a pair of pixels is left diagonal, for a 135 - degree scan.

### 3. Feature Extractions of Wood Defects

Based on gray level co-occurrence matrix, Haralick extracted 14 characteristics of images: energy  $E$  (also called the Angular Second Moment), Moment of inertia ( $I$ ) (also known as Contrast), Correlation coefficient  $C$  (Correlation), Variance  $VAR$  (Variance), deficit of IDM (Inverse Difference Moment), SOA, and the Average (Sum of business), and Variance  $SOV$  (Sum of Variance), Entropy  $SOE$  (Sum of Entropy), and Entropy  $H$  (Entropy), Difference Variance  $VOD$  (Variance of Difference), differential Entropy  $DOE$  (Difference of Entropy), information measure (including  $SOC$  (Shadow of Clustering) and  $POC$  (Prominence of Clustering)), the maximum system related to  $MP$  (Maximal aim-listed Probability). They reflect the gray distribution, information, and the texture feature of images from different angles.

After having tried many times, in this paper, the selection of step length is  $d = 2$ , for good recognition effect. Four commonly used statistics:

$$P_i = \sum_{j=1}^{Ng} P(i, j) \quad (2)$$

$$P_j = \sum_{i=1}^{Ng} P(i, j) \quad (3)$$

$$P_{i+j}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{Ng} \sum_{j=1}^{Ng} P(i, j) k \in \{2, 3, \dots, Ng\} \quad (4)$$

$$P_{i-j}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{Ng} \sum_{j=1}^{Ng} P(i, j) k \in \{0, 1, \dots, Ng - 1\} \quad (5)$$

Energy  $E$  :

$$E = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} [P(i, j)]^2 \quad (6)$$

Moment of inertia  $I$  :

$$I = \sum_{k=0}^{Ng-1} k^2 P_{i-j}(k) \quad (7)$$

Correlation  $C$  :

$$C = \frac{\sum_{j=1}^{Ng} \sum_{i=1}^{Ng} i \times j \times P(i, j) - \mu_1 \mu_2}{\delta_1 \delta_2} \quad (8)$$

which,  $\mu_1, \mu_2$  and,  $\delta_1, \delta_2$  are the Expectations, Standard Deviations of  $P_i, P_j$  ;

Variance  $VAR$  :

$$VAR = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (i - u)^2 P(i, j) \quad (9)$$

$u$  is the Expectation of matrix  $P$

Deficit moment  $IDM$  :

$$IDM = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{P(i, j)}{(1 + (i - j)^2)} \quad (10)$$

Sum average *SOA* :

$$SOA = \sum_{k=2}^{2Ng} k \times P_{i+j}(k) \quad (11)$$

Sum variance *SOV* :

$$SOV = \sum_{k=2}^{2Ng} (k - SOA)^2 \times P_{i+j}(k) \quad (12)$$

Entropy *H* :

$$H = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} P(i, j) \log(P(i, j)) \quad (13)$$

When  $P(i, j) = 0$ ,  $P(i, j) \log(P(i, j))$  gets the limit  $\lim_{x \rightarrow 0^+} (x \log x) = 0$ , the same conditions

in the same way below;

Sum Entropy *SOE* :

$$SOE = - \sum_{k=2}^{2Ng} P_{i+j}(k) \log(P_{i+j}(k)) \quad (14)$$

Differential variance *VOD* :

$$VOD = \sum_{k=0}^{Ng-1} (k - \sum_{k=0}^{Ng-1} k \times P_{i-j}(k))^2 P_{i-j}(k) \quad (15)$$

Differential Entropy *DOE* :

$$DOE = - \sum_{k=0}^{Ng-1} P_{i-j}(k) \log(P_{i-j}(k)) \quad (16)$$

Relevant information measure:

$$SOC = \frac{H - HXY1}{\max(H_1, H_2)} \quad (17)$$

$$POC = (1 - \exp(-2.0(HXY2 - H)))^{\frac{1}{2}} \quad (18)$$

$$\text{Which, } H_1 = - \sum_{k=2}^{2Ng} P_{i+j}(k) \log(P_{i+j}(k)), \quad H_2 = - \sum_{k=0}^{Ng-1} P_{i-j}(k) \log(P_{i-j}(k)),$$

$$HXY1 = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} P(i, j) \log(P_i P_j), \quad HXY2 = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} P_i P_j \log(P_i P_j) \quad (19)$$

The largest related system *MP* :

$$Q(i, j) = \sum_{n=1}^{Ng} \frac{P(i, n) P(j, n)}{P_i P_j} \quad (20)$$

*MP* took the second largest eigenvalue of matrix  $Q$ . In this article, if the denominator was 0 for calculating part of elements in matrix  $Q$ , that was hard to measure. So we deleted these eigenvalues, and only the top 13 eigenvalues in matrix were extracted and calculated.

In 26 training samples (No. 1 ~ 26, 1 ~ 15 are crack defect sample images, 16 ~ 23 sample image for knot flaws, 24 and 26 are decay defect sample images), some sample characteristic values are shown in Table 1.

**Table 1. Some Eigenvalues of Wood Defect Images**

Sample	E	I	C	VAR	IDM	SOA	SOV	SOE	H	VOD	DOE	SOC	POC
1	0.0129	11.92	0.998	3375.32	0.519	96.42	1420.38	3.752	5.732	9.78	1.667	-0.547	0.993
2	0.0290	1.99	0.981	51.11	0.606	81.95	202.21	3.474	4.203	1.40	1.130	-0.464	0.960
3	0.0028	12.39	0.990	621.39	0.362	152.92	2690.28	3.659	6.557	7.40	2.066	-0.469	0.991
4	0.0075	6.33	0.993	471.52	0.510	86.11	1192.91	4.430	5.759	4.71	1.535	-0.506	0.990
5	0.0107	3.14	0.991	172.19	0.548	77.10	685.14	4.281	5.367	2.10	1.354	-0.507	0.987
6	0.0115	14.09	0.995	1314.61	0.517	102.24	1301.03	4.003	5.533	12.04	1.607	-0.506	0.988
16	0.0030	3.98	0.993	283.72	0.472	122.98	1139.18	4.831	6.168	2.25	1.561	-0.512	0.993
17	0.0030	3.70	0.994	298.06	0.489	108.79	1199.78	4.890	6.168	2.15	1.516	-0.531	0.994
18	0.0002	184.53	0.941	1562.20	0.143	88.58	5962.25	2.771	8.845	99.24	3.298	-0.252	0.955
19	0.0067	3.74	0.993	256.28	0.515	157.28	2702.25	3.248	5.490	2.43	1.463	-0.529	0.990
24	0.0091	15.13	0.994	1189.50	0.513	46.73	2219.66	4.502	6.041	12.44	1.733	-0.517	0.992
25	0.0084	16.72	0.992	993.05	0.489	46.45	1882.67	4.483	6.088	13.40	1.832	-0.492	0.990
26	0.0269	20.05	0.977	434.48	0.601	38.77	1296.59	4.243	5.309	18.75	1.353	-0.563	0.992

In order to eliminate the factors out of image defects (Including scanning angle  $\alpha$ , pixel interval d, gray scale Ng A) that could influence the eigenvalues calculated by the gray level co-occurrence matrix, we use the "\*" form scanning to obtain the eigenvalues of the four directions and take the average value as the eigenvalue to eliminate the influence of scanning angle  $\alpha$ . And, we establish the recognition network by taking different pixel interval d and gray scale Ng, select the best recognition network corresponding interval d and gray scale Ng as the generation parameters of image gray level co-occurrence matrix. The eigenvalues, which are calculated by normalization gray level co-occurrence matrix of image, reflect the characteristics of the image in different ways and contain the image information of almost all aspects. The characteristic vector combined by characteristic values can effectively reflect the characteristics of the defect images.

#### 4. FBP Fuzzy Neural Networks

BP neural network is one of the widely applied feed-forward neural network models, as shown in the Fig. 2. Within the same nerve layer, between each neuron are independent, without connection. Between two adjacent nerve layer neurons, input and output are connected (linear relationship) by weight (undetermined constants). By using smooth transfer function as the activation of neurons function, they make response (activation treatment, in general is a nonlinear operation) to the input signal (incentives), and combine the response with weight combination to make it as the input

signal of next layer neurons. General BP neural networks have one or more (generally no more than 2) hidden layers.

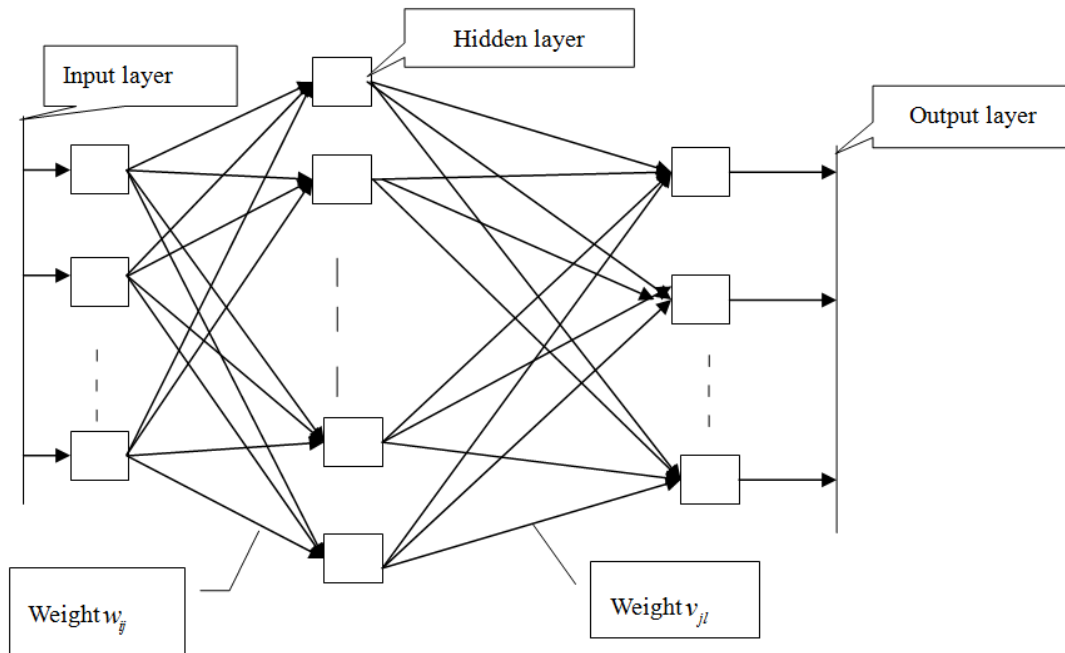


Figure 2. BP Neural Networks

putting  $X = (x_1, x_2, \dots, x_p)^T$  in input layers, via the linear combination of  $w_{kj}$   $k = 1, 2, \dots, p$  and threshold  $\theta_k$ .

$$v_k = \sum_{j=1}^p w_{kj} x_j - \theta_k \quad (21)$$

As the Number  $k$  hidden layer of the input signal, the incentive of transfer function  $\varphi(\bullet)$ , and get the response output

$$y_k = \varphi(v_k) \quad (22)$$

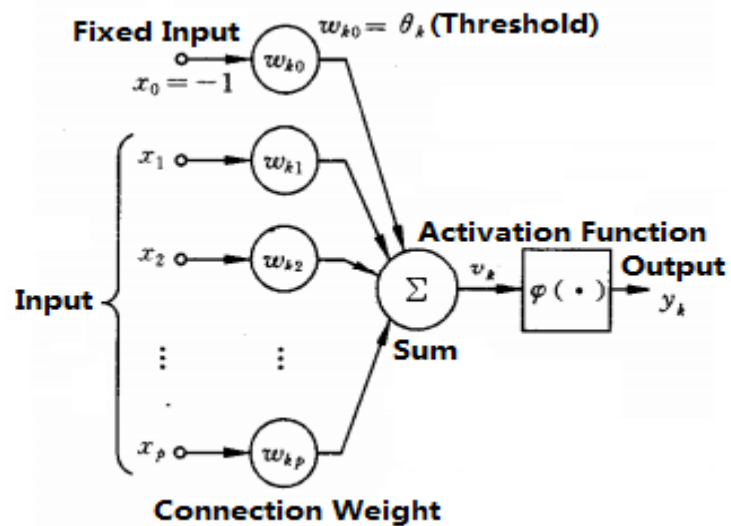


Figure 3. The BP Network from Input Layer to Hidden Layer

If adding an input layer neurons, the corresponding increase one dimension of input vector  $x_0 = -1$ , the corresponding weight  $w_{k0} = \theta_k$ .

$$\text{Then, } v_k = \sum_{j=0}^p w_{kj} \cdot x_j \quad (23)$$

So the thresholds can be regarded as weights.

The parameters in the BP neural network are obtained through learning and training the network [11]. Learning and training process of BP neural networks is as follows: the input signal goes from the input layer through the hidden layer to output layer to get the network output (model forward-transmission in networks); calculating the error signals between actual output and desired output in networks (learning the trained goals), and then correcting parameters layer by layer from the output layer through the middle layer to the input layer (errors backward-propagation in networks) [12]. Model forward-transmission and errors backward-propagation, in turn, have alternating repeated (network learning and training). When errors are less than the expected target, learning and training of networks get end and the output in networks has saved.

Fuzzy neural network (FNN) brings together the advantages of the fuzzy theory and neural network, and becomes the collection of recognition, association, learning and adaption of fuzzy information. J.J.Buckly *et al* firstly proved that the FNN can approximate any continuous fuzzy function for any precision [13], and then many scholars have proved that the fuzzy neural network has global approximation ability. In this paper, we improved the traditional BP neural network based on the inspiration of recursive fuzzy neural network system[14], introduced fuzzy mathematics into the hidden layer of BP network. We changed the transfer function in hidden layer, made the form of input-output like ladder shape, expanded the general two states 0 and 1 into interval [0, 1], and modified the BP neural network into fuzzy neural network. That is, we designed a neural network with fuzzy hidden layers [15], shorthand for FBP network.

FBP consists of input layer, hidden layer and output layer. The input layer contains  $n$  neurons,  $n$  dimensions of the feature vector of samples, and  $N$  characteristic vectors totally. The  $k$ -th feature vector is  $\mathbf{x}^k = (x_1^k, x_2^k, \dots, x_n^k)^T, k=1,2,\dots,N$ . Output layer contains  $m$  neurons,  $m$  is the number of defect types,  $\mathbf{o}^k = (o_1^k, o_2^k, \dots, o_m^k)^T$  and  $\mathbf{r}^k = (r_1^k, r_2^k, \dots, r_m^k)^T$  are the actual output and desired output for  $X^k$  in FBP network respectively.

The hidden layer designs. Hidden layer has one layer in all. The number of hidden layer neurons corresponds to the requirement of questions and the amount of input and output unit numbers. To design the number of nodes for the pattern recognition/classification may refer to the following formula.

$$q = \sqrt{N + m + a} \quad (24)$$

Which,  $q$  is the number of hidden layer neurons,  $N$  is the number of input layer neurons,  $m$  is the number of output layer neurons;  $a$  is the constant range from 1 to 10. In this paper,  $N=13$ ,  $m=3$ , and taking  $a=9$  via trying many times, for better network identification effect.

## 5. FBP Network Learning and Training

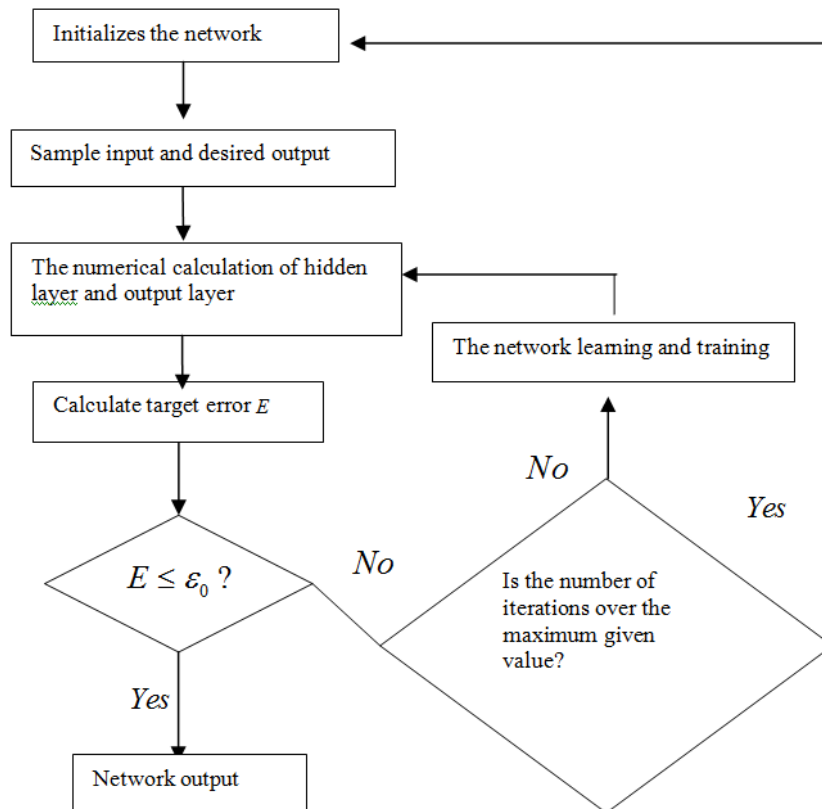
This paper uses the BFGS quasi-newton algorithm and the L - M algorithm in matlab8.0 (64-bit) study platform for training of the network, the results after many attempts in Table 2. FBP network flow chart as shown in Fig. 4.

By contrast, although the effect of L - M algorithm is better than quasi-newton algorithm for the learning and training of FBP network, when the new sample data have

entered network the test result of L - M algorithm is clearly worse than quasi-newton algorithm. Namely, in this situation, L - M algorithm for FBP appears the phenomenon of excessive study. So we select the quasi-newton algorithm as the training learning algorithm for FBP. After many attempts, we select the best effect of FBP network as the final network for recognition.

**Table 2. Quasi-newton Algorithm and L - M Algorithm Network**

Training samples			The test results of new samples			
Training algorithm	Training objectives (MSE)	Time-taken	Practical training error (MMSE)	Percentage of successful recognition for training samples	MSE	Average recognition success rate
quasi-newton	0.0002	4~5 Sec.	A little less than 0.0002	100%	About $10^{-2}$	Above 90%
L - M		About 2 Sec.	About $10^{-6}$	100%	$>0.5$	80%~90%



**Fig. 4 FBP Flow Chart**



## 6. Conclusion

In this paper, fuzzy mathematics has organically combined with the BP neural network to build FBP network. The feature vectors of samples are the input for FBP network, and the FBP network is learned and trained by supervised Hebb learning-rate and BFGS quasi-newton algorithm. When the network gets the goal of expected errors, we stop the network learning and training, and save the network at this time. The average recognition rate of FBP network is more than 90% for the test samples, the effect of the linear regression analysis between network output membership degree and expectation membership degree for test samples is ideal.

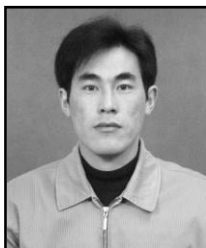
## Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities (No. 2572014CB30), the Project of Natural Science Foundation of Heilongjiang Province of China (No.C201338), the Project of National Science Foundation of China (No.31170518), the Project of Science and technology research of department of education of Heilongjiang province (No.12543019) and Science and technology innovative talents special funds of Harbin (No.2012RFXXG080).

## References

- [1] J. Li, "Wood Science", Higher Education Press, China, (2002).
- [2] D. W. Qi, "A study on image processing system of non-destructive log detection", *Scientia Silvae Sinica*, vol. 92-96, (2001).
- [3] X. D. Zhang, W. Sun and J. L. Jin, "Multilayer feedforward fuzzy neural network for image recognition", *Computer Applications and Software*, vol. 5, (2001), pp. 1-4.
- [4] H. B. Mu, "Wood defect detection based on BP and RBF neural network research", Harbin: the Northeast Forestry University Engineering Department, (2010).
- [5] B. Kosko, "Fuzzy associative memories", Addison Wesley, (1987).
- [6] J. J. Li, A. X. Wang and G. Y. Zhang, "Pattern recognition", Electronic industry press, (2006).
- [7] D. Y. Wu, N. Yi and X. Q. Su, "Wood defect recognition based on gray level co-occurrence matrix and cluster method", *Computer and Digital Engineering*, vol. 38, (2010), pp. 38-41.
- [8] B. Andrea and P. Flavio, "An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 33, no. 2, (2003).
- [9] L. A. Zadeh, "Fuzzy sets", *Inform. and Control*, vol. 8, (1965), pp. 338-353.
- [10] E. B., "Probabilite'et certitude", Press Universitaires de France, (1950); Paris, France.
- [11] P. Y. Liu, "Fuzzy associative memory based on a modular arithmetic of fuzzy neural network", *Journal of Systems Engineering and Electronics*, vol.19, (1997), pp. 54-58.
- [12] Q. G. Yang, J. Zhang and S. Z. Zhang, "Forward neural network learning algorithm based on the quasi-newton method", *Control and Decision*, vol. 12, (1997), pp. 357-360.
- [13] J. J. Buckley and Y. Hayashi, "Can fuzzy neural nets approximate continuous fuzzy functions", *Fuzzy Sets Systems*, vol. 61, (1994), pp. 43-51.
- [14] S. T. Wang, "Fuzzy system, Fuzzy neural network and the application design", Shanghai science and technology literature publishing house, (1998).
- [15] X. Z. Zeng, "X-ray real-time imaging detection technology field measurement method", *Journal of NDT*, vol. 13, no. 5, (2000), pp. 27-32.

## Authors



### Hongbo Mu

Author's profile. Place of Birth: TongLiao, China

Date of Birth: 07-24-1980

E-mail: mhb-506@163.com

Mobile: +8613796610867

Associate professor, Northeast Forestry University, Harbin, China

