Research of simulation techniques based on rough set theory

Weiliang Zhu*

Qinggong College, Hebei United University, Tangshan, Hebei 063000, China

Received 1 March 2014, www.tsi.lv

Abstract

Rough set theory can effectively analyse and deal with incomplete information in simulation techniques. This paper studied the knowledge reduction problem and discrete continuous attributes and improved the BP neural network on the basis of rough set theory. Firstly, methods of attribute reduction of classical are analysis. This paper proposes a heuristic algorithm for reduction of knowledge based on information entropy. Subsequently, it studied the combination algorithm of rough set and neural network. Pre-treatment of sample data based on rough set in dealing with imprecision and uncertainty issues on the edge. The decision rules obtained after reduction in order to map to the training sample of neural network. Finally, the neuron number of hidden layer of neural network and hidden layer makes the neural network more logical. The simulation results show that the simulation technique of rough set and neural network has obvious complementary and reduce the time to train the neural network. It improved the training accuracy and generalization ability simulation techniques achieved satisfactory results.

Keywords: simulation techniques, rough set, knowledge reduction, BP neural network

1. Introduction

The computer simulation technology is the use of computer science and technology building is a comprehensive technology simulation model system and dynamic experiment of the model under some experimental conditions. Human thinking is the organic combination of image thinking and logical thinking. And the integration of rough set and neural network is not accidental, but rather reflect the qualitative and quantitative, human intelligence is clear and implied and serial and parallel cross the conventional thinking mechanism. Intelligent hybrid system is some intelligent technology to combine the use to overcome the shortcomings of single species technology, realize the complementary advantages. So as to facilitate the processing of information in different type in one integrated system. Combining rough set and neural network is the research focus in the field of intelligent information processing. Knowledge acquisition, knowledge expression and reasoning decision rules from a large number of observations and experimental data is an important task for intelligent information processing system. Especially for the inaccurate, incomplete knowledge, rough set theory and artificial neural network shows the infinite charm, rough set method to simulate the human abstract logical thinking, intuitive thinking simulation image of the neural network method.

The literature [1, 2] presented a neural network based on rough sets theory, which consists of rough neurons and conventional neurons. Rough neuron consists of a traditional neuron, the data in the upper and lower boundaries as input or output values of the network. According to the literature of [3] neural network PID controller general difficult to obtain defect prediction system of output value, we put forward an improved rough set neural network PID controller. The literature [4-6] based on neural network put forward an evaluation method of rough set and neural network project, the method uses data to enrich the data of rough set theory. From the given learning found a set of rules in the sample, extraction rules as the input of neural network. This method simplifies the structure of the neural network, improve the training efficiency. Literature [7-9] presented a neural network model based on rough sets, the method using the data analysis method of rough set, extracted from the data input output mapping rules to the subspace. Analysis and modelling on the stability of rock slope. And compared with the traditional modelling method of neural network, illustrates the effectiveness of the method. The literature [10-13] and put forward a new method of discrete interval of a division of property, according to this classification, a method is proposed for reduction and the establishment of a joint pattern recognition system based on rough sets and LVQ neural network.

The computer simulation technology research starting from the rough set model and elaborates the basic theory of rough set. Study on the discretization method of continuous attributes reduction of knowledge. In view of the limitation of the rough set model, the extended model of fuzzy rough set model and variable precision rough set model are discussed in detail. Based on rough set theory and neural network in the processing logic of complementarily, the combining algorithm application in

*** *Corresponding author*** e-mail: 915799718@qq.com

system modelling, and the corresponding simulation, and achieved good results in this paper.

2 Rough set theory

Rough set cannot confirm individuals vest in border areas, and the boundary area is defined as the upper approximation set and lower approximation set difference set. Because it has certain mathematical formula description, complete by data decision, so as to avoid the influence of subjective factors. In data pre-treatment process, the rough set theory can be applied to feature more accurate extraction. Data preparation process, using rough set theory, data reduction properties of data set for dimension reduction operation. In data mining stage, the rough set theory is used for classification rule discovery. In the interpretation and evaluation process, the rough set theory can be used for the result of statistical evaluation. The basic method of rough set has been in decision-making, forecasting, uncertainty reasoning, the network planning, and the ensemble.

A relational database can be viewed as an information system, the column to property and the line to object. Suppose $P \subset A$, $x_i, x_j \in U$, definition two binary relation $IND(P)$ called equivalence relation:

$$IND(P) = \{(x_i, x_j) \in U \times U \mid \forall p \in P, p(x_i) = p(x_j)\}$$

If and only if $P(x_i) = P(x_j)$ for all $p \in P$, then x_i, x_j is equivalence relation about the property set P.

Suppose $X \subset U$ is a subset of the individual domain in the information system $S = \{U, A, V, f\}$, then the lower approximation and upper approximation set and the boundary region of X are as follows:

$$\underline{P}X = \{Y \in U / P : Y \subseteq X\},$$

$$\overline{P}X = \{Y \in U / P : Y \cap X \neq \varnothing\},$$

$$Bnd_P(X) = \overline{P}X - \underline{P}X,$$

where $\underline{P}X$ are elements in $X \subset U$ and must be classified collection, that is the maximum of definable sets in X and $\overline{P}X$ are elements in U and must be classified collection, that is the minimum of definable sets in X. $BndP(X)$ is the set of elements which cannot be classified in $X \subset U$.

Lower and upper approximation schema: assume that there is an information system, there are two attributes. Attributes of a 5 value, attribute two has 6 values. For example, the upper approximation and lower approximate relation as follows:

TABLE 1 Upper approximation and lower approximate relation

U	Headache	Temp.	Flu
U_1	Yes	Normal	No
U_2	Yes	High	Yes
U_3	Yes	Very-high	Yes
U_5	No	High	No
U_6	No	Very-high	Yes
U_7	No	High	Yes
U_8	No	Very-high	No

Suppose $S = \{U, A, V, f\}$ is a information system and $X \subset U$, $P \subseteq A$, then the accuracy of approximation of X in S is as follows:

$$\mu_P(X) = \frac{\underline{\mu}_P(X)}{\overline{\mu}_P(X)} = \frac{card(\underline{P}X)}{card(\overline{P}X)}.$$

Note: $card(X)$ is the number of elements in set X.

Suppose S is a information system and $P \subseteq A$. Let $\psi = \{x_1, x_2, \ldots, x_n\}$ is a classification of U and $X_i \subseteq U$, then the $P-$ lower approximations and the $P-$ upper approximations of ψ can expressed as follows:

$$\underline{P}\Psi = \{\underline{P}X_1, \underline{P}X_2, \ldots, \underline{P}X_n\},$$

$$\overline{P}\Psi = \{\overline{P}X_1, \overline{P}X_2, \ldots, \overline{P}X_n\}.$$

The quality of classification of ψ, which conformed by the attribute subset $P \subseteq A$ is as follows:

$$\gamma_P(\psi) = \frac{\sum_{i=1}^{n} card(\underline{P}X_i)}{card(U)}.$$

The classification quality represented by the ratio of all object attribute subset P of correct classification number and the number of information system.

3 BP neural network

Neural network also can deal with inaccurate and incomplete knowledge. However rough sets theory and artificial neural network method of two different, rough set method to simulate human abstract thinking, neural network method is used to simulate the image intuitive thinking. Neural network general cannot handle with semantic form of input rough set theory can input qualitative and quantitative or mixed information neural network can be realized without guide.

Sum up the biological neurons in the process of transfer information biological neurons is a multiple input and single output unit. When the nerve cell j has many inputs x_i(i=1,2,…,m) and single output y_j, then the relationship between input and output can be expressed as follows:

$$\begin{cases} s_j = \sum_{i=1}^{m} w_{ij} x_i - \theta_j \\ y_j = f(s_j) \end{cases} .$$

The w_{ij} is connection weights from neurons i to neurons j, and f is the transfer function or called excitation function. After determining the neuron model, topology structure and learning method of characteristics and the ability of a neural depends on the network.

BP Neural network is error back propagation neural network is referred to as ", it consists of an input layer, one or more of the hidden layer and an output layer structure, each time by a certain number of neurons composition. These neurons as a man of nerve cells are interrelated. The structure as shown in Figure 1 shows:

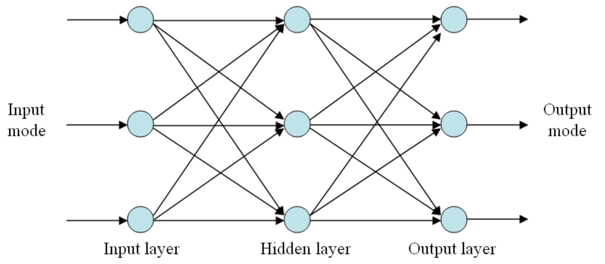

FIGURE 1 The BP neural network model

The traditional rough set of lack of semantic, rough neurons into semantic structure. Rough neural consists of a pair of overlapping normal neuron: $\overline{r}$ (Upper Neural) and $\underline{r}$ (Lower Neural), neurons output $output\,\overline{r}$ is always greater than the lower neuron output $output\,\underline{r}$:

$$output\,\overline{r} = \max(f(input\overline{r}), f(input\underline{r})) ,$$

$$output\,\overline{r} = \min(f(input\overline{r}), f(input\underline{r})) .$$

The function f is the transfer function.

4 Algorithm design

This template provides authors with most of the formatting specifications needed for preparing their articles.

Neural network can deal with imprecise and incomplete knowledge. However, the rough set method both theory and artificial neural network simulation method is different, the abstract thinking of human rough set, neural network method simulation image intuitive thinking. It has different characteristics.

The basic BP algorithm includes two aspects: the counter-propagating signals prior to the dissemination and error. The calculation of actual output from input to output direction, while the weights and thresholds of correction from the output to the input direction.

4.1 FORWARD PROPAGATION PROCESS SIGNAL

The input net_i of the first i node in hidden layer as follows:

$$net_i = \sum_{j=1}^{M} w_{ij} x_j + \theta_i .$$

The output y_i of the first i node in hidden layer as follows:

$$y_i = \phi(net_i) = \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right).$$

The input net_k of the first k node in hidden layer as follows:

$$net_k = \sum_{i=1}^{q} w_{ki} y_i + a_k = \sum_{i=1}^{q} w_{ki} \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right) + a_k .$$

The output o_k of the first k node in hidden layer as follows:

$$o_k = \psi(net_k) = \psi\left(\sum_{i=1}^{q} w_{ki} y_i + a_k\right) = \psi\left(\sum_{i=1}^{q} w_{ki} \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_i\right) + a_k\right) .$$

4.2 BACK PROPAGATION OF ERROR

Back propagation error, namely first by the output layer to output error calculation of each layer of neurons, and then to adjust the weights of each layer and the threshold error gradient descent method, so that the final output of the modified network can meet its expected value.

For the two type of error criterion, function E_p of each sample p as follows:

$$E_p = \frac{1}{2} \sum_{k=1}^{L} (T_k - o_k)^2 .$$

The total system error criterion function on the training sample is as follows:

$$E = \frac{1}{2} \sum_{p=1}^{P} \sum_{k=1}^{L} (T_k^p - o_k^p)^2 .$$

We can get the correction output layer threshold Δa_k and the correction of hidden layer weights Δw_{ij} and the correction of hidden layer threshold $\Delta\theta_i$ according to the error gradient descent method are modified correction output layer weights Δw_{ki}. The output layer weights adjustment formula as follows:

$$\Delta w_{ki} = -\eta \frac{\partial E}{\partial w_{ki}} = -\eta \frac{\partial E}{\partial net_k} \frac{\partial net_k}{\partial w_{ki}} = -\eta \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial net_k} \frac{\partial net_k}{\partial w_{ki}}$$

And formulas as follows:

$$\frac{\partial E}{\partial y_i} = -\sum_{p=1}^{P}\sum_{k=1}^{L}(T_k^p - o_k^p)\cdot\psi'(net_k)\cdot w_{ki}$$

So finally obtained the following formula:

$$\Delta w_{ki} = \eta\sum_{p=1}^{P}\sum_{k=1}^{L}(T_k^p - o_k^p)\cdot\psi'(net_k)\cdot w_{ki}\cdot\phi'(net_i)\cdot x_j$$

$$\Delta\theta_i = \eta\sum_{p=1}^{P}\sum_{k=1}^{L}(T_k^p - o_k^p)\cdot\psi'(net_k)\cdot w_{ki}\cdot\phi'(net_i)$$

Additional momentum method with the help of network in the correction of the weight, not only consider the role of error in the gradient, and considering the changing trends in the error surface. In the absence of additional momentum effect, the network may get into local minimum shallow value, using the additional momentum effect may be over these minima. The method is every weight in based on back-propagation method on changes on a proportional plus on previous weight or threshold changes in the amount of value, and to generate new weights based on the back propagation method or threshold changes. With additional momentum factor weights and threshold adjustment formula:

$$\Delta w_{ij}(k+1) = (1-mc)\eta\delta_i p_j + mc\Delta w_{ij}(k),$$

where k is the number of training and MC is the momentum factor and generally around 0.95.

Usually adjust the learning rate criterion is: check whether the real weight reduces the error function, and if it does, then the learning rate is small, can increase an amount; if not, the overshoot, so would reduce the value of the learning rate. The formula gives an adaptive learning rate adjustment formula:

$$\eta(k+1) = \begin{cases} 1.05\eta(k) & if\ E(k+1) < E(k) \\ 0.7\eta(k) & if\ E(k+1) > 1.04E(k) \\ \eta(k) & \text{other} \end{cases},$$

where $E(k)$ is the first k step error sum of square. The initial learning rate of the selected range has great arbitrariness.

The calculated properties of mutual information in the greedy algorithm to process it is difficult to judge the importance of the breakpoints. Then, consider the example of those breakpoints can be separated from the most. With the next breakpoint mutual information to a breakpoint to chose which one breakpoint when a breakpoint importance of the same. The improved algorithms are as follows:

Input: decision information system $S = (U, A, d)$.

Output: a discrete interval set CUT.

Step 1: according to the information system S construction of the new information system S^*.

Step 2: the initial breakpoint $CUT = \Phi$.

Step 3: in each column of S^*, calculation of the number 1 in the last row.

Step 4: the breakpoint to join in the nuclear CUT, $S_0 = S^*$.

Step 5: breakpoint at will add to place CUT in the columns of S_0 and on this column value is 1 line, get information table S_1.

Step 6: if S_1 is empty, then turn to step 13.

Step 7: Re-calculate the number of column 1 in the last line of S_1.

Step 8: in the last row if only the maximum value of a column of S_1, then the row where the breakpoint is added to the CUT, $S_0 = S_1$, turn to step 5.

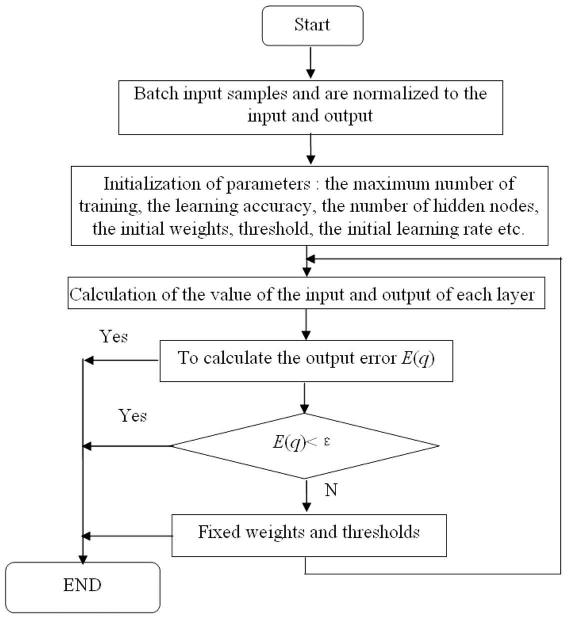

FIGURE 2 The flowchart of the improved BP algorithm program

Step 9: set the maximum value of the column has a maximum value n, respectively, where the column and the column value is 1 line, get n information tables $S_{11}, S_{12}, ..., S_{1n}$.

Step 10: if there is any empty table in $S_{11}, S_{12}, ..., S_{1n}$, have to break a table space to join in CUT, turn to step 13.

Step 11: in each column of $S_{11}, S_{12}, ..., S_{1n}$, calculation of the number 1 in the last row. Were calculated for each table in the last line of column values corresponding to the maximum of the breakpoint of S_{1i} for mutual information in the table S are the breakpoint.

Step 12: take the minimum mutual information such as not only a column is the one corresponding to the breakpoint and get the breakpoint S_{1i} to join in *CUT*, $S_0 = S_{1i}$, turn to step 9.

Step 13: the end.

5 Simulation procedure

The simulation results show that, the design method of the control system performance improved significantly. Information system decision table as shown in the table:

TABLE 2 The decision table information system

U	a	b	d
1	0.8	2.0	1
2	1.0	0.5	0
3	1.3	3.0	0
4	1.4	1.0	1
5	1.4	2.0	0
6	1.6	3.0	1
7	1.3	1.0	1

TABLE 3 S^* of Table 2

U^*	p_1^a	p_2^a	p_3^a	p_4^a	p_1^b	p_2^b	p_3^b	SUM
(u_1,u_2)	1	0	0	0	1	0	0	3
(u_1,u_3)	1	1	0	0	0	0	1	3
(u_1,u_5)	1	1	1	0	0	0	0	3
(u_2,u_4)	0	1	1	0	1	0	0	3
(u_2,u_6)	0	1	1	1	1	1	1	6
(u_2,u_7)	0	1	0	0	1	0	0	2
(u_3,u_4)	0	0	1	0	0	1	1	3
(u_3,u_6)	0	0	1	1	0	0	0	2
(u_3,u_7)	0	0	0	0	0	1	1	2
(u_4,u_5)	0	0	0	0	0	1	0	1
(u_5,u_6)	0	0	0	1	0	0	1	2
(u_5,u_7)	0	0	1	0	0	1	0	2
SUM	3	5	6	3	4	6	5	

It exist number 1 identical at the column of p_3^a and p_2^b is the largest in Table 3. At this time were the maximum value of columns and the value is 1 line.

TABLE 4 Information table S_{11}

U^*	p_1^a	p_2^a	p_4^a	p_1^b	p_2^b	p_3^b
(u_1,u_2)	1	0	0	1	0	0
(u_1,u_3)	1	1	0	0	0	1
(u_2,u_7)	0	1	0	1	0	0
(u_3,u_7)	0	0	0	0	1	1
(u_4,u_5)	0	0	0	0	1	0
(u_5,u_6)	0	0	1	0	0	1
SUM	2	2	1	2	3	3

TABLE 5 Information table S_{12}

U^*	p_1^a	p_2^a	p_3^a	p_4^a	p_1^b	p_3^b
(u_1,u_3)	1	1	0	0	0	1
(u_1,u_5)	1	1	1	0	0	0
(u_2,u_4)	0	1	1	0	1	0
(u_2,u_7)	0	1	0	0	1	0
(u_3,u_6)	0	0	1	1	0	0
(u_5,u_6)	0	0	0	1	0	1
SUM	2	4	3	2	2	2

Table 4 shows that the column of p_3^a and then remove column value is 1 rows. We calculate the mutual information in the table S for the breakpoint p_3^a as follows:

$$I(p_3^a,p_2^b)=H(p_2^b)-H(p_2^b\,|\,p_3^a)=$$
$$-(\frac{3}{7}\log_2\frac{3}{7}+\frac{4}{7}\log_2\frac{4}{7})-[\frac{3}{7}(\frac{1}{3}\log_2\frac{1}{3}+\frac{2}{3}\log_2\frac{2}{3})+$$
$$\frac{4}{7}(\frac{2}{4}\log_2\frac{2}{4}+\frac{2}{4}\log_2\frac{2}{4})]=1.9502$$

$$I(p_3^a,p_3^b)=H(p_3^b)-H(p_3^b\,|\,p_3^a)=$$
$$=-(\frac{5}{7}\log_2\frac{5}{7}+\frac{2}{7}\log_2\frac{2}{7})-[\frac{3}{7}(\frac{1}{3}\log_2\frac{1}{3}+\frac{2}{3}\log_2\frac{2}{3})+$$
$$\frac{4}{7}(\frac{2}{4}\log_2\frac{2}{4}+\frac{2}{4}\log_2\frac{2}{4})]=1.8282$$

Table 5 shows that the column of p_2^b and then remove column value is 1 rows, the p_2^a column number 1 to a maximum of 4. Calculation of the mutual information in the table S for the breakpoint p_2^b as follows:

$$I(p_2^b,p_2^a)=H(p_2^a)-H(p_2^a\,|\,p_2^b)=$$
$$=-(\frac{5}{7}\log_2\frac{5}{7}+\frac{2}{7}\log_2\frac{2}{7})-[\frac{3}{7}(\frac{1}{3}\log_2\frac{1}{3}+\frac{2}{3}\log_2\frac{2}{3})+$$
$$\frac{4}{7}(\frac{1}{4}\log_2\frac{1}{4}+\frac{3}{4}\log_2\frac{3}{4})]=1.7203$$

Because breakpoint p_2^a on the breakpoint p_2^b of minimum mutual information, thus remove the breakpoint is p_2^a and p_2^b, the two breakpoint columns and on this column value is 1 of the line, and the information table is S_1. In fact, because of the breakpoint p_2^b is to distinguish objects 4, 5 only a breakpoint in this case the algorithm, and add the first breakpoint CUT and remove the breakpoint, the information table S_1 is shown as follows:

TABLE 6 Information table S_1

U^*	p_1^a	p_2^a	p_3^a	p_4^a	p_1^b	p_3^b
(u_1,u_3)	1	1	0	0	0	1
(u_1,u_5)	1	1	1	0	0	0
(u_2,u_4)	0	1	1	0	1	0
(u_2,u_7)	0	1	0	0	1	0
(u_3,u_6)	0	0	1	1	0	0
(u_5,u_6)	0	0	0	1	0	1
SUM	2	4	3	2	2	2

From the Table 6, the column p_2^a number 1 to a maximum of 4. The breakpoint columns and on this column values for the 1 line, the table is in Table 5. The following process is no longer here.

This algorithm integrates the advantages of several algorithms. Because of the amount of information mutual information represents a source obtained from another source size, so take small mutual information column as a breakpoint makes its own more information. Satisfactory effect was obtained through the example. It solved the limitation of former algorithms effectively.

6 Conclusions

The computer simulation technology has the advantages of high efficiency, safety, environmental conditions of constraint, can change the time scale etc. It has become an important tool for analysis, design, operation, evaluation, and training of complex systems. Neurons in the hidden layer of neural network and the hidden layer number in order to map to the training sample of neural network. To make neural network more logical, and reduce the time to train the neural network, improve the training accuracy and generalization ability. Based on the research of related literature, the rough set and weak coupling mode using the automatic control in the neural network system, the training sample pre-treatment of the neural network, improve the training speed. In pattern recognition, it is using the concept of rough membership function, constructing sub neural network pattern recognition system based on rough set. The output of the simulation computing improved the anti-noise ability and the accuracy of pattern recognition. It reduced the fuzzy neural network training time, and improving the precision of training.

References

[1] Zdzislaw Pawlak, Andrzej Skowron 2007 Rough sets and Boolean reasoning *Information Sciences Publishers* Zdzislaw Pawlak chapter 2

[2] HUANG Xian-xiang, LONG Yong 2010 Summarization of Distributed Visual Simulation Technology *Journal of System Simulation*. **22**(11) 201-7

[3] Jian Meng, Minghui Cui, Kaiquan Shi 2006 Variation rough sets in the application of set pair analysis *Journal of Shandong University* **36**(3) 21-7

[4] Mantas C J, Puche J M, Mantas J M 2006 Extraction of similarity based fuzzy rules from artificial neural networks *Journal of Approximate Reasoning* **43**(2) 111-7

[5] Ganivada A, Dutta S, Pal S 2011 Fuzzy rough granular neural networks, fuzzy granules, and classification *Journal of Theoretical Computer Science* **41**(4) 53-61

[6] Kwak K C 2011 A Development of Cascade Granular Neural Networks *Journal of IEICE Transactions on Information and Systems* **94**(7) 307-15

[7] Fernandez Navarro F, Hervas-Martinez C, Sanchez-Monedero J 2011 MELM-GRBF A modified version of the extreme learning machine for generalized radial basis function neural networks *Journal of Neural Computing* **74**(16) 116-27

[8] Martinez J M, Escandell-Montero P Soria-Olivas E 2011 Regularized extreme learning machine for regression problems *Journal of Neural Computing* **74**(17) 71-7

[9] Chong Wang, Li Zhao Intermediate 2010 View Synthesis Based on Adaptive BP Algorithm and View Interpolation *JCIT* **10**(5) 1068-71

[10] Chen H Y, Leou J J 2012 Saliency-directed color image interpolation using artificial neural network and particle swarm optimization *Journal of Visual Communication and Image Representation* **23**(1) 35-41

[11] Zhu Bing, Wang Hong-fang, Wang Wen-Sheng, Li Yao-Qing 2007 Analysis of relation between flood peak and volume based on set pair *Journal of Sichuan University* **39**(3) 121-30

[12] Gao J, Gao X 2008 A New Model for Credibilitic Option Pricing *Journal of Uncertain systems* **4**(2) 33-41

[13] C Robert, Wang L F, Alam M 2012 Training neural networks using Central Force Optimization and Particle Swarm Optimization: Insights and comparisons *Journal of Expert Systems with Applications* **1**(39) 25-37

[14] Mohammed A, Minhas R, Wu Q M J 2011 Human face recognition based on multi dimensional PCA and extreme learning machine *Journal of Pattern Recognition* **44**(101) 55-67

[15] Mohamed M H 2011 Rules extraction from constructively trained neural networks based on genetic algorithms *Journal of Neuronal computing* **17**(74) 112-19

[16] Kahramanli H, Allahverdi N 2009 Rule extraction from trained adaptive neural networks using artificial immune systems *Journal of Expert Systems with Applications* **2**(36) 153-61

Authors

Weiliang Zhu, born in June20, 1979, China

Current position, grades: assistant professor, teacher of Qinggong College of Hebei United University.
University studies: mathematical model of information processing and application.
Scientific interest: Adaptive control, Rough set, Research on Intelligent Algorithm
Publications: 3 (EI)
Experience: The main author of the improved adaptive robust control problem