A Method of Remote Fault Diagnosis Based on Multilayer SOM

CHEN Wen-yu WANG Xiao-bin Sun Shi-xin
School of Computer Science and Engineering
University of Electronic Science & Technology of China
Chengdu Sichuan , China
cwy@uestc.edu.cn

LIU Jing-bo
School of Mathematics and Computer Science
Chongqing Three Gorges University
Chongqing , China
allahljb@163.com

***Abstract*—This paper establishes the model of remote fault detect system, using typical fault data of a type of plane's undercarriage braking system to implement the diagnosis. The result shows that he Ensemble Method of Remote Fault Diagnosis based on PCA and ICA and Multi-layer SOM, integrates the advantage of PCA in low noise and ICA in high noise. Multilayer SOM can classify the input data rapidly with associative memory and parallel processing. The computing results accord with the actual situation and indicate that the ensemble PCA-ICA-MSOM method has large potential in recognizing remote patterns.**

***Keywords*—Remote fault diagnosis, PCA, ICA, multilayer SOM, ensemble.**

I. INTRODUCTION

Remote faults are always highly paroxysmal and complex. Remote Fault Diagnosis ensemble system is a pattern recognition process on machine's state essentially[1]. The Rule-Based Expert System, Artificial Intelligence System, and Neural Networks System are most common methods in Remote Fault Diagnosis ensemble system. It is effective when the Rule-Based Expert system deals with shallow knowledge, but inefficient on complex systems consisting of deep knowledge. Self-Organizing Feature Map (SOM) is based on the basic tenets of the brain structure, which means cells that have similar functions congregate. SOM simulates these characteristics to make output processing unit similar to its neighbors after network study. Therefore, SOM could be used to establish model of the Remote Fault Diagnosis ensemble system. The model is self-learning and its activation process is a one-way feed- forward manner, rather than many-times feedback manner, so it has fast reasoning speed. This paper would take PCA、ICA、MSOM methods to construct a Remote Fault Diagnosis ensemble system.

(1)To remove noise, the fault data can be processed by mean.

(2)The pattern recognition sub-system is based on PCA-MSOM, ICA-MSOM.

(3) The ensemble technique of the subsystem's pattern recognition can improve the correct rate of faults diagnosis effectively BASIC PRINCIPLES AND METHODS

A. Remove Noise by Mean

Detection data $x(t)$ is superposed by valid data $z(t)$ and noise $g(t)$.

$$x(t) = z(t) + g(t) \tag{1}$$

Noise signal is originated by multiple noise sources. Each noise source has no decisive effect, so the noise signal observes the normal distribution, and can be zero-mean. Using mean-value method to remove noise data can increase the correct rate of diagnosis. If $E(g(t)) = 0$, then $E(x(t)) = E(z(t)) + E(g(t)) = E(z(t))$ .valid data:

$$z(t) = \frac{1}{k}\sum_{i=1}^{k} x(t-\mathrm{i}) \tag{2}$$

B. Principal Component Analysis

In multidimensional space problem research, principal component analysis,(PCA for short), because the variables usually have many dimensions, and variables of different dimensions always have relevance between each other, and this relevance causes redundant information of the detection data. PCA is a typical statistical analyze method to find out what are "the major components of data. Using PCA can reduce dimensionality when loss data only contain a small amount of information[2]. PCA is often used for data compression and characteristic extraction.

1) Eigenvalue and Characteristic vector Method: The correlation matrix $x x^T$ of vector group x is a symmetric matrix. Find out the characteristic roots of the correlation matrix and sort them in descending order: $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$;find out the characteristic vectors corresponding to every characteristic root, $e_1, e_2, \cdots, e_n$,let $w = (e_1, e_2, \cdots e_m)$. Suppose y_i is the i^{th} major component, then $y_i = e_i^T x$ ($i = 1, 2 \cdots m$). Choose the former m major components, the number m is depended by the designated contribution rate g (usually, $g \ge 0.99$). Contribution rate g is defined as following: $g = (\sum_{i=1}^{m} \lambda_i)/(\sum_{j=1}^{n} \lambda_j)$ the matrix expression of the major component is:

$$y = w^T x \tag{3}$$

2) Neural Network Method [3]: Let matrix $x = (x_1, x_2, \cdots, x_k)$, where, $x_1, x_2, \cdots, x_k$ are n-dimensional input column vectors, let

978-1-4244-1674-5/08 /$25.00 ©2008 IEEE CIS 2008

matrix $y=(y_1,y_2,\cdots,y_k)$ where, $y_1,y_2,\cdots,y_k$ are m-dimensional output column vectors. In which :

$$x_i=(x_{1,i},x_{2,i},\cdots,x_{n,i})^T \quad (i=1,2,\cdots,k)$$

$$y_i=(y_{1,i},y_{2,i},\cdots,y_{n,i})^T \quad (i=1,2,\cdots,k)$$

Suppose $w_{i,j}$ is the weight of the link between input neuron i and output neuron j . $w_i=(w_{i,1},w_{i,2},\cdots w_{i,n})^T$, $(i=1,2,\cdots,m)$.
Weighted matrix $w=(w_1,w_2,\cdots w_m)^T$ .The output equation of the neural network is:

$$y=w\cdot x \quad (4)$$

3) Algorithm for weight learning in neural network: The initial value of $w_j(1)$ ($j=1,2,\cdots m$) is chosen randomly, and

$$w_j(k+1)=w_j(k)+\mu\cdot y_j(k)[x(k)-\sum_{s=1}^{j}w_s(k)\cdot y_s(k)] \quad (5)$$

Continue leaning until $|w_j(k+1)-w_j(k)|$ is less than the designated threshold.

C. Independent Component Analysis

Independent component analysis (ICA for short) is an analyze method based on the signals' higher-order statistics to deal with blind signal separation. After the decomposition of ICA, the signal components are mutual independent[4]. ICA extracts important part of the data so as to maximally identify the entirety[5]. The principal is: based on known observed signal v and unknown transformation matrix A , we need to find out the unknown source signal s , that is $s=B\cdot M\cdot v$, in the formula, $v=A_{m\times n}\cdot s$, $x=M_{n\times m}\cdot v$, $s=B_{n\times n}\cdot x$, $s(t)=(s_1(t), s_2(t),\cdots,s_k(t))$ are k n-dimensional data source signals at time t , $v(t)=(v_1(t), v_2(t),\cdots,v_k(t))$ are k m-dimensional observed values of the sensor at time t , $n\geq m$.

The steps of FastICA:

(1) Perform the whitening process to the data If vector x satisfies the condition $\begin{cases}E\{x\}=0\\E\{xx^T\}=I\end{cases}$, then we say that x satisfies whitening process, we can use PCA method to perform the process to vector x .

(2) Generate vector randomly, where, $|w(1)|=1$.

(3)

$$w(k+1)=C^{-1}\cdot E\{x(w(k)^T x)^3\}-3w(k) \quad (6)$$

In the formula, $C^{-1}=E(x\cdot x^T)$ is the covariance of vector x .

(4)To normalize $w(k+1)$

$$w(k+1)^*=\frac{w(k+1)}{|w(k+1)|} \quad (7)$$

(5)If $|w^T(k+1)\cdot w(k)|$ is not close enough to 1, Then let $k=k+1$, and return to step(3). Otherwise, output the matrix equation: $s=\mathrm{w}^{\mathrm{T}}\cdot v$.

D. Self-Organizing Feature Map

Self-organizing feature map (S0M) is also known as Kohonen network. It was proposed by a Finnish scholar named Teuvo Kohonen in 1981. The network is a self-organizing and self-learning network composed by fully connected neuron arrays[6]. The idea is: neurons distributed within the space are demultiplexed. When a neural network accepts input mode from external world, it will divide into different regions of reaction. Different regions have different reacting features on input mode.

Multilayer self-organizing feature map (MSOM for short) creates the maps between multidimensional point sets to two-dimensional point sets[7]. Two-dimensional space is also the mapping of a special multi-dimensional space that can proceed to sets of two-dimensional space points. Several SOM networks can be combined hierarchically together as an MSOM. After the 2D process, the former SOM layer's output can be the input of next SOM layer. Except for the 1st SOM layer, other input layers are all two-dimensional..

The learning process of MSOM network:

(1) Initialize. Assigning a small weight to $w_{ij}(0)$, which is the weight of the link of input neuron and the output neuron. Assign the neighborhood function $S_j(0)$.

(2) Provide new input vector x ;

(3) Calculate the distance d_j between input sample and every output neuron j .

$$d_j=\|X-W_j\|=\sqrt{\sum_{i=1}^{N}(x_i(t)-w_{ij}(t))^2} \quad (8)$$

Find out the neuron k , thus for every j , the formula $d_k=\min_j(d_j)$ be satisfied.

(4) Provide the neighborhood function $S_k(t)$ which decreases as time increases.

(5) Adjust weight, suppose,

$$w_{ij}(t+1)=w_{ij}(t)+\eta(t)[x_i(t)-w_{ij}(t)] \quad (9)$$

Out of which, $\eta(t)=\frac{1}{t}$ or $\eta(t)=0.2(1-\frac{t}{10000})$

(6) Calculate the output O_k :

$$O_k=f(\min_j\|x-w_j\|) \quad (10)$$

In the formula, f function is 0-1 function or other nonlinear functions;

(7) If O_k reaches the total training sum, output O_k . Otherwise, return to step(2).

(8) If O_k is not output by the last layer of SOM, to 2D O_k ,and let O_k be the input of next SOM layer, return to step (1), go to the next SOM layer.

To 2D: If the SOM output layer outputs O_k , and the output layer is an $m\times n$ neuron array, to 2D the array to (i,j) .

$$i\times m+j=k \qquad 0\leq i<m, 0<j\leq n \quad (11)$$

II. USING MSOM NEURAL NETWORK TO IMPLEMENT THE REMOTE DIAGNOSIS OF THE PLANE'S UNDERCARRIAGE BRAKE SYSTEM

A. Ensemble remote diagnosis system based on MSOM neural system

The main idea of ensemble learning[8] is using multiple methods to solve the same question, keeping the advantage of each method and eliminating the disadvantage to get a better result. Each method may pre_treat the data in order to get a better result in pattern recognition. For example, ICA is combined with neural network and eural network which uses

PCA method is able to filtrate the noise signal in the data[9][10][11].

After experiments, we find that there are some features of using methods of PCA-MSOM and ICA-MSOM: in the condition of low noise, the MSOM processed by PCA has a higher accuracy than that processed by ICA; in the condition of high noise, the MSOM processed by ICA has a higher accuracy than that processed by PCA. Shown in the Figure.3.

Combine the ICA component extraction technology with the method of MSOM pattern recognition to construct a remote fault diagnosis system model. As shown in Figure 1.

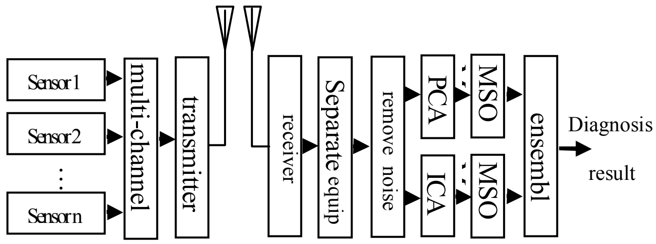

Figure 1. Structure model of Remote Fault Diagnosis system.

The model contains several main parts:

(1) Hardware. Including sensors, multi-channel equipment, transmitters, receivers and by-passing equipment.

(2) Data pretreatment. Including reduce the noise of the detection data; use eigenvalue and eigenvector for PCA principal component analysis; use FastICA method for ICA independent component analysis.

(3) Neural Network Self-Organizing Map. For the data analyzed by PCA,ICA ,construct an MSOM to self-organizing map learning .

(4) Integration of information . Integrate the pattern recognized by MSOM to get the final diagnosis resultIn the learning process, the major component and the independent component obtained via PCA and ICA methods are not unique, neither is the feature map pattern after MSOM feature mapping process. Therefore he neural network can not recognize the result directly after training. After training the neural network, simulate the typical faulty data to PCA-MSOM,ICA-MSOM subsystem to get mapping relation y.

$y: I_k \to O_j^{(i)}$, I_k Is the k th kind of typical faulty inputs $O_j^{(i)}$ $(i=1,2;\ j=1,2,\cdots,k)$ are k kinds of outputs in the two subsystems respectively, This is a one-to-one subsystems eversible mapping, its inverse mapping recorded as y^{-1}.

Suppose $\omega(i) = f(\sigma)$ and $\omega(i) \geq 0$, $(i=1,2)$ and $\omega(1)+\omega(2)=1$,they are the weights of the PCA-MSOM sub-classifier and ICA-MSOM sub-classifier, σ is noise variance. The output of the ensemble classify system O_m is defined as:

$$O_m = \mathrm{sgn}(\omega(1) \cdot y^{-1}(O_j^{(1)}) + \omega(2) \cdot y^{-1}(O_j^{(2)})) \qquad (12)$$

$$(m, j = 1, 2, 3, \cdots, k)$$

B. Experiment

Use this system to train and test typical fault samples of a type of plane's undercarriage braking system.(Table 1).

TABLE I. REPRESENTATIVE MODEL OF BRAKE FAULT

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
y_1	0.0	0.0	0.3	0.2	0.3	0.5	0.0	0.3	0.5
y_2	0.7	0.5	0.0	0.3	0.1	0.7	0.8	0.1	0.2
y_3	0.2	0.4	0.3	0.1	0.1	0.1	0.4	0.1	0.1
y_4	0.2	0.2	0.4	0.2	0.2	0.2	0.2	0.2	0.2
y_5	0.0	0.3	0.5	0.2	0.2	0.4	0.2	0.2	0.2
y_6	0.3	0.1	0.2	0.1	0.0	0.3	0.5	0.1	0.3
y_7	0.1	0.1	0.0	0.2	0.3	0.1	0.2	0.3	0.1
y_8	0.2	0.2	0.2	0.0	0.2	0.6	0.2	0.1	0.3
y_9	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.2
y_{10}	0.0	0.1	0.3	0.0	0.0	0.3	0.1	0.0	0.3
y_{11}	0.0	0.1	0.3	0.2	0.1	0.3	0.1	0.0	0.3
y_{12}	0.2	0.3	0.1	0.2	0.3	0.1	0.2	0.2	0.2

x is sign set $(1,2,\cdots,9)$, y is fault set $(j=1,2,3,\cdots,12)$

Adding noise signals with square obviation values as 0.01, 0.05, 0.1, 0.2, 0.4, into 12 types of typical faulty data. Every group of signal is a 9*1200 matrix and every fault has 100 testing data. The parameter of noise reduction k = 10; PCA contribution rate $g=0.99$; in MSOM each SOM layer learns 500 steps The adaptation function is trains, the initialize function is initlay and the training function is `trainr`. Dimension of input layer in first layer of SOM neural network is 9 and the input layer is an 8*8 neuron matrix. Dimension of input layer in second and third layers in SOM neural network is 2, and the output layer is a 6×6 neuron matrix. Let $\omega(1)=1$ $\sigma \leq 0.15$, $\omega(1)=0$ $\sigma > 0.15$. Calculate the correct recognition rate of the noise-reduced detection data fault pattern, calculate 20 times to get the average of correct recognition rate.

For single layer SOM, only PCA process the 1, 2, 3 layer SOM, only ICA process the 1, 2, 3 layer SOM process the remote fault Pattern Recognition. Abscissa is the noise variance, longitudinal is fault correct recognition rate, recognition rate is shown in Figure 2.

Using integrated system to recognize the remote fault pattern, normal single SOM, SOM only process PCA and ICA, their comparison on the remote fault pattern recognition rate is shown in Figure3,Figure4.

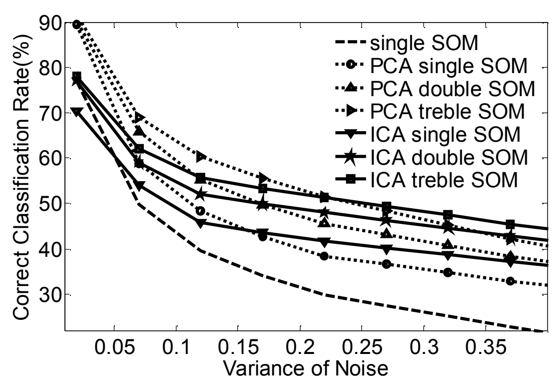

Figure 2. Comparison of correct classification rates.

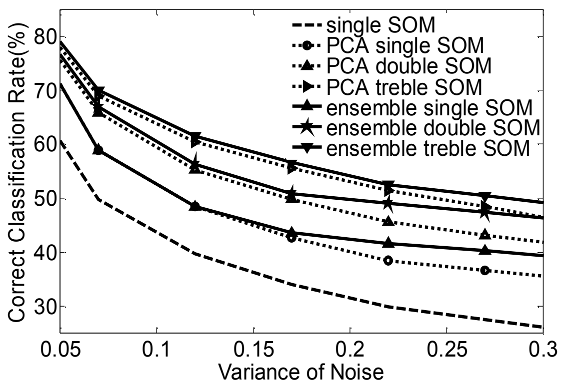

Figure 3. Comparison of ensemble and PCA-MSOM.

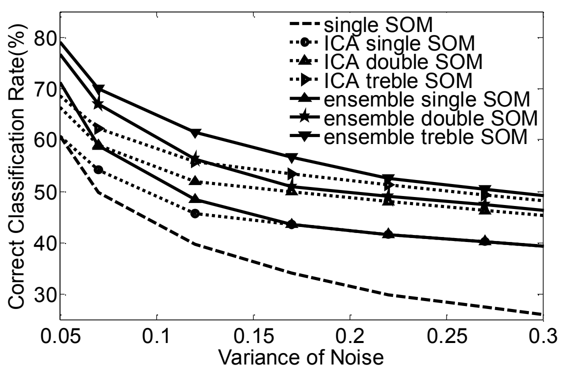

Figure 4. Comparison of ensemble and ICA-MSOM correct classification rates

Results show that:

(1)The correct rate of faults diagnosis is improved as the number of SOM layers increase However, the increasing rate decreases.

(2) In the condition of high noise, the correct rate of faults diagnosis with only ICA is higher than that with only PCA. On the contrast in the condition of low noise, the correct rate of faults diagnosis with only PCA is higher than that with only ICA.

(3) The correct rate of faults diagnosis with ICA or PCA is higher than that with normal SOM.

(4) The Ensemble Method of Remote Fault Diagnosis based on PCA and ICA and Multi-layer SOM is better than normal SOM, PCA-MSOM and ICA-MSOM.

III. SUMMARY

This paper proposes a Remote Fault Diagnosis ensemble system, which is integrated by two MSOM(Multilayer Self-Organizing Feature Map) subsystems. In order to remove noise, the fault data can be processed by Mean-value then the two subsystems process the features extracted by PCA and ICA respectively. Finally, the ensemble system combines the advantages of the two subsystems. In experiments, the proposed methods have been successfully evaluated using twelve different datasets. The result shows that he Ensemble Method of Remote Fault Diagnosis based on PCA and ICA and Multi-layer SOM, integrates the advantage of PCA in low noise and ICA in high noise. Multilayer SOM can classify the input data rapidly with associative memory and parallel processing.

References

[1] Qu Liang-sheng, Zhang Hai-jun, Some Basic Problems in Machinery Diagnostics, China mechanical engineering, Vol.11, No.1-2P.211-216, 2000.

[2] E. Oja, A simplified neuron model as a principal component analyzer[J]. Math.Biol. Vol.15, pp267-273, 1982.

[3] Zhang yi, Wang Ping-an, Zhou Ming-tian ,Computing Eigenvalues and Eigenvectors of Matrix by Neural Networks[J], Chinese uournal of computers,23(1), 71-76,2000.

[4] OJA E. From Neural Learning to Independent Components [J], Neuro Computing, 1998, (22): 187-199.

[5] Jongsun Kim, Jongmoo Choi, Juneho Yi, Matthew . Using ICA for Face Recognition Robust to Local Distortion and Partial Occlusion. IEEE Transactions on Pattern Analysis and Machine Intelligenc-e, Vol. 27, No. 12, 2005.12

[6] KohonenT. The Self-Organizing Map, Proceedings of IEEE,1990,78(9):1464-1481.

[7] Lei Wei, Fei Zhong, Tao He ,etc. Fault Pattern Recognition Based on Unsupervised Neural Network [J]. Computer Measurement & Control ,2006,14(6):742-744

[8] Zhou Kaili, Kang Yaohong.Study on Performance for Pattern Recognition Systems Based on Neural Network Data Fusion [J], Computer Engineering,2006.9 , Vol32,No17.

[9] Wang Yan, Chen Xiang-dong, Jiang Ya-dong, etc. Study on electronic-nose recognitionmethod with independent component analysis and back-propagation neural networ[J]. Transducer and Micro-system Technologies,2006,25(9):36-39

[10] Lu Shuang, Hou Yue-qian. FaultPattern Recognition of Rolling Bearing Based on Principal Components Analysis and Radial Basis Function Neural Network[J]. Maching Tool & Hydraulics, 2005,3:185-188

[11] Sven Nordholm ,Siow Yong Low. Speech Signal Extraction Utilizing PCA-ICA Algorithm whith Non-Uniform Spacing Microphong Array[J].ICASSP 2006 V 965-968