INTELLIGENT DIAGNOSIS METHOD FOR PLANT MACHINERY USING WAVELET TRANSFORM, ROUGH SETS AND NEURAL NETWORK

PENG CHEN[1], TAKAYOSHI YAMAMOTO[2], TETUROU MITOMA[1], ZHONG-YONG PAN[1], XIN-YING LIAN[1]

[1] Department of Environmental Science & Technology, Mie University, 1515 Kamihama, Tsu, Mie, Japan.
[2] Project of Machinery Diagnosis, Japan Science and Technology Agency
E-MAIL: chen@bio.mie-u.ac.jp; yamamoto@extranet.jp

Abstract:

This paper proposes an intelligent diagnosis method for plant machinery using wavelet transform (WT), rough sets (RS) and partially-linearized neural network (PNN) to detect faults and distinguish fault type at an early stage. The WT is used to extract feature signal of each machine state from measured vibration signal for high-accurate diagnosis of states. The decision method of optimum frequency area for the extraction of feature signal is discussed using real plant data. We also propose the diagnosis method by using "Partially-linearized Neural Network (PNN)" by which the type of faults can be automatically distinguished on the basis of the probability distributions of symptom parameters. The symptom parameters are non-dimensional parameters which reflect the characteristics of time signal measured for condition diagnosis of plant machinery. The knowledge for the PNN learning can be acquired by using the Rough Sets (RS) of the symptom parameters. The practical examples of diagnosis for rotating machinery are shown to verify the efficiency of the method.

Keywords:

Condition diagnosis; vibration signal; wavelet transformation; rough sets; neural network

1. Introduction

In the field of machinery diagnosis, particularly rotating machinery diagnosis, vibration signals (displacement, velocity and acceleration) are measured for detection of failure and discrimination of failure types. When the vibration signals for the diagnosis are measured at an early stage of the machine failure or at a location remote from the failed parts, the extraction of symptom parameters and discrimination of failure types are difficult, because the power of noise is much stronger than that of the actual failure signal. Therefore, it is important that the noise be isolated from the measured signal as far as possible, in order to be able to sensitively identify the failure types[1].

Furthermore, in the case of condition diagnosis of plant machinery, the knowledge for distinguishing failures is ambiguous because definite relationships between symptoms and fault types cannot be easily identified. The main reasons can be explained as follows. (1) It is difficult to identify the symptom parameters for diagnosis by which all fault types can be distinguished perfectly. (2) In the early stages of a fault, effects of noise are so strong that the symptoms of a fault are not evident.

The conventional neural network (NN) cannot reflect the possibility of ambiguous diagnosis problem. Furthermore, the NN will never converge, when the symptom parameters put in the 1st layer have the same values in different states [2].

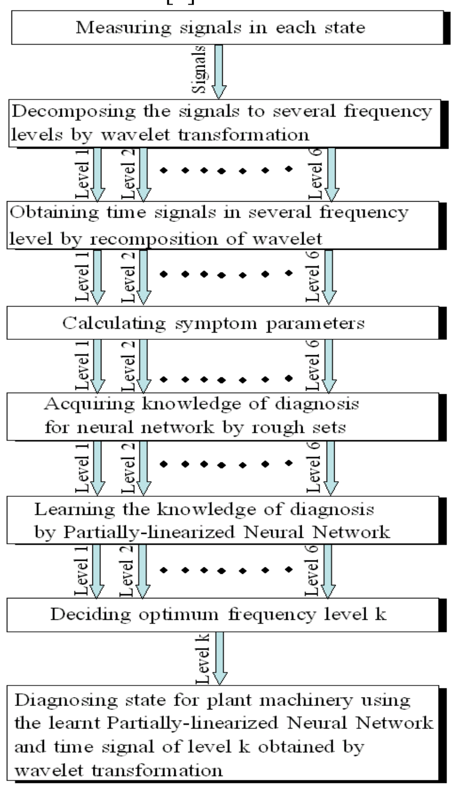

Figure. 1 Flowchart of the processing for this study

For above reasons, we propose an intelligent diagnosis method for plant machinery using wavelet transform (WT), rough sets (RS) and partially-linearized neural network (PNN) to detect faults and distinguish fault type at an early stage. Fig.1 shows the flowchart of the processing for this study.

The noise cancellation for extraction of feature signal in an optimum frequency area is carried out by wavelet transformation. The data (the diagnosis knowledge) for the PNN learning are acquired by using the Rough Sets [4]. In this paper, practical examples of fault diagnosis of rotating machinery will verify that the method is effective.

2. Extraction of feature signal by wavelet transformation

The wavelet transform of time signal $f(t)$ is expressed by

$$(W_{\psi}f)(b,a)=\frac{1}{\sqrt{a}}\int_R f(t)\overline{\psi(\frac{t-b}{a})}dt \qquad (1)$$

In this paper, the object of diagnosis is the structural faults, such as unbalance, misalignment and looseness etc., that are often occurring in a shaft of rotating machinery, and these faults may cause serious machine accidents and bring great production losses.

Because the feature signals of these faults are appear in low frequency area, we use the ReverseBior analyzing wavelet $\psi(t)$ [3] to extract the feature signals. Fig. 2 shows the example of ReverseBior 1.5.

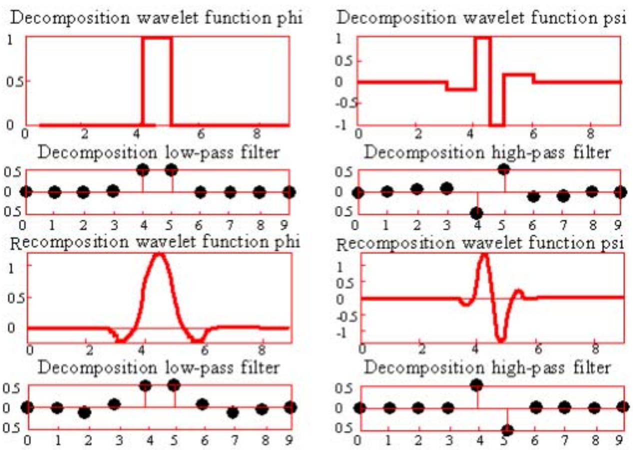

Figure.2 Example of ReverseBior 1.5.

Fig.3 shows the rotating machine, the shaft and the accelerometer for the diagnosis. Fig. 4 shows vibration signals measured in each state. There are noises with strong power in the measured raw signals. The sampling frequency for the measurement is 5.12kHz. The states to be diagnosed for the rotating machine are normal state, misalignment, unbalance and looseness.

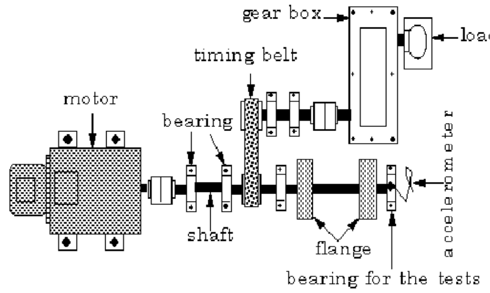

Figure. 3 Rotating machine for diagnosing

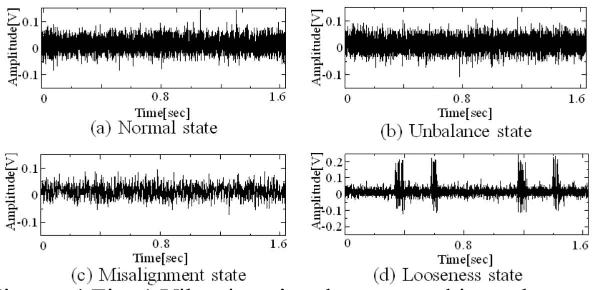

Figure. 4 Fig. 1 Vibration signals measured in each state

We use the ReverseBior analyzing wavelet to decompose the signals to 6 levels in low frequency as shown in Fig. 5.

Level	Freq. area
Raw	0~5.12kHz
1	0~2.56kHz
2	0~1.28kHz
3	0~640Hz
4	0~32Hz
5	0~16Hz
6	0~8Hz

Figure.5 Frequency area of each level

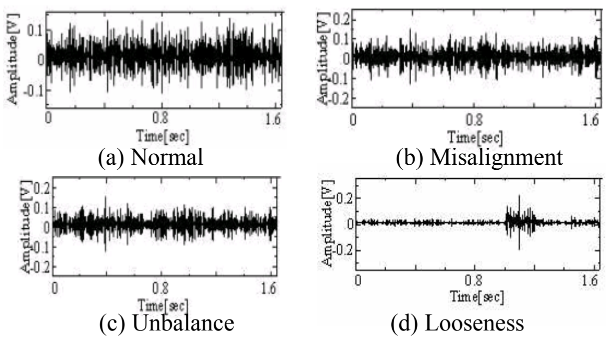

Figure. 6 Time signal in level 4 form 0Hz to 32Hz

Fig. 6 shows the recomposed time signals in each state. In next paragraph, we shall show that optimum frequency area is the level 4 from 0Hz to 32Hz for diagnosing these states.

3. Partially-Linearized neural network

The neuron numbers of m-th layer of a NN is N_m. The set $X^{(1)}=\{X_i^{(1,j)}\}$ expresses the pattern inputted to the 1st layer and the set $X^{(M)}=\{X_i^{(M,k)}\}$ is the teacher data to the last layer (M-th layer). Here, $i=1\ to\ P,\ j=1\ to\ N_1$, $k=1\ to\ N_M$, and,

$X_i^{(1,j)}$: The value inputted to the j-th neuron in the input (1st) layer;

$X_i^{(M,k)}$: The output value of k-th neuron in the output (M-th) layer; $k=1\ to\ N_M$

Even if the NN converge by learning $X^{(1)}$ and $X^{(M)}$, it cannot well deal with the ambiguous relationship between new $X^{(1)*}$ and $X_i^{(M)*}$, which have not been learnt. In order to predict $X_i^{(M)*}$ according to the probability distribution of $X^{(1)*}$, partially linear interpolation of the NN is introduced as Fig. 7, we called it "Partially-linearized Neural Network (PNN)".

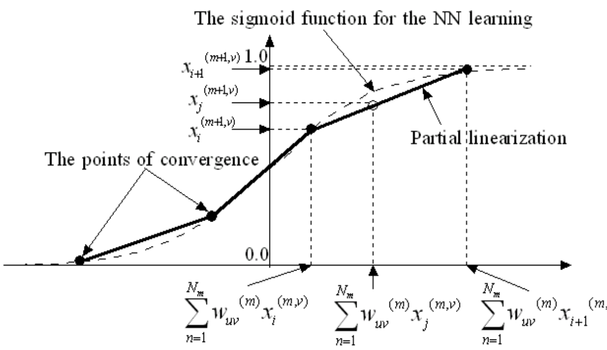

Figure. 7 The partial linearization of the sigmoid function

In the NN which has converged by the data $X^{(1)}$ and $X^{(M)}$, the symbols are used as follows.

$X_i^{(m,t)}$: The value of t-th neuron in the hidden (m-th) layer; $t=1\ to\ N_m$

$W_{uv}^{(m)}$: The weight between the u-th neuron in the m-th layer and the v-th neuron in the ($m+1$)-th layer; $m=1\ to\ M; u=1\ to\ N_m; v=1\ to\ N_{m+1}$

If these values are all remembered by computer, when new values $X_j^{(1,u)*}$ ($X_j^{(1,u)}<X_j^{(1,u)*}<X_{j+1}^{(1,u)}$) are inputted to the first layer, the predicted value of v-th neuron ($v=1$ to N_m) in the ($m+1$)-th layer ($m=1$ to $M-1$) will estimated by

$$X_j^{(m+1,v)}=X_{i+1}^{(m+1,v)}-\frac{\{\sum_{\mu=0}^{Nm}W_{uv}^{(m)}(X_{i+1}^{(m,u)}-X_j^{(m,u)})\}(X_{i+1}^{(m+1,v)}-X_j^{(m+1,v)})}{\sum_{\mu=0}^{Nm}W_{uv}^{(m)}(X_{i+1}^{(m,u)}-X_i^{(m,u)})} \quad (2)$$

In the above way, the sigmoid function is partially linearized as shown in Fig. 7. If a function, such as Fig. 8, need to be learnt, the PNN will learn the points ● shown in Fig. 8. When new data (s_1', s_2') are inputted into the converged PNN, the value symbolized by ■ correspond to the data (s_1', s_2') will be quickly identified as P_e shown in Fig. 8. So the PNN can be used to deal with ambiguous diagnosis problems.

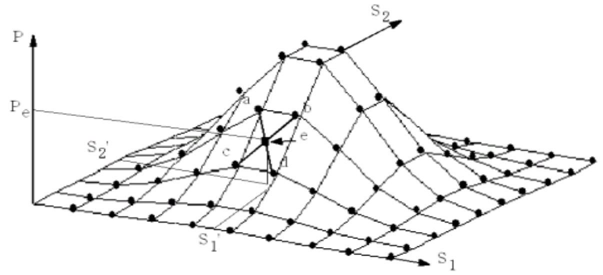

Figure. 8 Interpolation by the PNN

4. Knowledge acquisitions by rough sets

In order to acquire the diagnosis knowledge by rough sets [4], the data x_j must be measured in each state. The total number of the data x_j is J. The values of symptom parameters ${}^jp_{1s}\cdots{}^jp_{ms}$ can be calculated by vibration signal measured in each state. The ${}^jp_{is}$ must be digitized as the teacher data for the PNN by the following formula.

$${}^jp_{is}=0\ to\ A_{pi}=\mathrm{int}[{}^jp_{is}/\{(\max\{{}^jp_{is}\}-\min\{{}^jp_{is}\})/N_{pi}\}+1] \quad (3)$$

Here, int[x] is the function which gives the integral values of x.

$$P=\{p_1,p_2,\cdots,p_m\} \quad (4)$$

is the initial symptom parameter set. ${}^jP_S$ is the set of symptom parameter values measured in the state S.

$${}^jP_S=\{{}^jp_{1S},{}^jp_{2S},\cdots,{}^jp_{mS}\} \quad (5)$$

$[xj]_{Ri}$ and A set are shown as follows :

$$[x_j]_{Ri}=r_i=\{x_j \mid x_j\quad and\quad x_t\in[x_j]_{Ri}\rightarrow{}^jp_{i\bullet}={}^tp_{i\bullet};j,t=1\ to\ J\} \quad (6)$$

$$A=\{r_i \mid i=1\ to\ Q\}=\{[x_j]_{Ri}\} \quad (7)$$

The value sets (r_{iP} and r_{jP}) of symptom parameters of r_i and r_j are :

$${}^{ri}P=\{{}^{ri}p_{1.},{}^{ri}p_{2.},\cdots,{}^{ri}p_{m.}\};\quad {}^{rj}P=\{{}^{rj}p_{1.},{}^{rj}p_{2.},\cdots,{}^{rj}p_{j.}\} \quad (8)$$

The symptom parameters set P_{ij} selected from P shown in formula (4), which can discriminate between r_i and r_j, is :

$$P_{ij} = \{p_k \big| p_k \in P;\ p_k^* \textit{ is the value of } P_k;\ p_k^* \in {}^{ri}P \\ \textit{or } p_k^* \in {}^{ri}P \rightarrow P_k^* \in ({}^{ri}P \cup {}^{rj}P) - ({}^{ri}P \cup {}^{rj}P)\} \tag{9}$$

For distinguishing r_i(i=1 to Q) from r_j (j=1 to Q, $j \neq i$), there may be redundant symptom parameters in the initial set P shown in formula (4). In order to remove the redundant symptom parameters, the following algorithm is proposed.

(a) Removing p_i from P;

(b) Calculating P_{ij} shown in formula (9).

(c) If $P_{ij} \neq \Phi$ (empty set), then p_i is the redundant symptom parameter. Removing p_i from P. Returning to (a) and repeating from (a) to (c) and from $i=1$ to $i=Q$;

(d) After removing all of redundant symptom parameters, the new set of symptom parameters $P' = \{p_1, p_2, \cdots, p_l\} \quad (l \leq m)$ is obtained and the values set of P' of r_i is :

$${}^{ri}P' = \{{}^{ri}p_1, {}^{ri}p_2, \cdots, {}^{ri}p_l\} \tag{10}$$

The possibility ${}^{S}\beta_{ri}$ of state S expressed by r_i can be calculated by

$${}^{s}\beta_{ri} = \frac{card({}^{s}x_j)}{card(x_i)}\% \tag{11}$$

Here, card(·) is the element number. ${}^{S}x_j \in r_i$ is the x_j obtained from state S.

According to the above principle, the input data and teacher data (diagnosis knowledge) for PNN are as follows.

Input data : The value sets ${}^{ri}P'$ (formula (10)) of symptom parameters of r_i (i=1 to Q), from which redundant symptom parameters have been removed.

Teacher data : The possibility ${}^{S}\beta_r$ (formula (11)) of state S.

5. Symptom parameters for condition diagnosis

The symptom parameters in the set of formula (4) used for condition diagnosis are the non-dimensional symptom parameters in time domain. The x_i is digital data of vibration signal.

$$p_1 = \frac{\sigma}{\overline{X}};\ \text{Here, } \overline{X} = \frac{\sum_{i=1}^{N} |x_i|}{N} \tag{12}$$

$$p_2 = \frac{\sum_{i=1}^{N} (|x_i| - \overline{X})^3}{\sigma^3};\quad \text{Here, } \sigma = \sqrt{\frac{\sum_{i=1}^{N} (|x_i| - \overline{X})^2}{N-1}} \tag{13}$$

$$p_3 = \frac{\sum_{i=1}^{N} (|x_i| - \overline{X})^4}{\sigma^4} \tag{14}$$

$$p_4 = \overline{X}_p / \overline{X} \tag{15}$$

Here, $\overline{X}_p$ is the mean value of peak values of $|x_i|$ (i=1 to N).

$$p_5 = \overline{|X_{\max}|} / \overline{X}_p \tag{16}$$

Here, $\overline{|X_{\max}|}$ is the mean value of 10 peak values (from maximum peak value to tenth value).

$$p_6 = \overline{X}_p / \sigma_p \tag{17}$$

Here, σ_p is the standard deviation of peak values of $|x_i|$ (i=1 to N).

$$p_7 = \overline{X}_L / \sigma_L \tag{18}$$

Here, x_L and σ_L are the mean value and the standard deviation of valley values of x_i respectively.

$$p_8 = \frac{\sum_{i=1}^{N} \sqrt{|x_i|}}{\sqrt{\sigma}} \tag{19}$$

$$p_9 = \frac{\sum_{i=1}^{N} x_i^2}{\sigma^2} \tag{20}$$

$$p_{10} = \frac{\sum_{i=1}^{N} \log(|x_i| + 1)}{\log(\sigma)};\quad x_i \neq 0 \tag{21}$$

$$p_{11} = \frac{\sum_{i=1}^{N} \exp(|x_i| + 1)}{\exp(\sigma)} \tag{22}$$

6. Example of diagnosis and verification

The values of the symptom parameters shown form formula (12) to (22) are calculated by using the time signals of vibration in each level obtained by wavelet transformation, and they are digitized as the teacher data for the PNN by formula (3). The redundant symptom parameters are removed by the algorithm shown form step (a) to step (d) in last paragraph. Fig. 9 shows redundant symptom parameters in each level by the mark "X". For example, we can distinguish each state by only using p_1, p_5 and p_8 in level 4.

Symptom parameters →	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}
Raw	O	X	X	O	X	X	O	X	X	X	X
Level 1	O	X	X	X	O	O	X	X	X	O	O
Level 2	X	X	X	X	O	X	X	O	X	X	X
Level 3	X	X	O	X	O	X	O	X	O	X	X
Level 4	X	X	x	X	X	O	X	X	X	O	X
Level 5	X	X	O	O	X	X	X	X	X	O	X
Level 6	O	O	X	X	X	X	X	X	X	X	X

Figure. 9 redundant symptom parameters in each level

Fig. 10 shows the parts of the acquired knowledge of diagnosis for PNN learning by rough sets in each level. Here, ${}^{N}\beta_{ri}$, ${}^{M}\beta_{ri}$, ${}^{U}\beta_{ri}$ *and* ${}^{L}\beta_{ri}$ are the possibility grades of normal state, misalignment, unbalance and looseness states respectively. The result of comparison of detection rate in each state with each frequency level is shown in Fig. 11. By this figure, it is obvious that the detection rate of level 4 for distinguishing the states is higher than that of other levels. So we used the recomposed time signals by wavelet transformation to diagnose the states.

p_1	p_4	p_7	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
3	4	4	0.25	0	0.75	0
8	1	1	0	1	0	0
3	4	4	0	0	1	0
4	7	7	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

(a) Data using raw time signal

p_1	p_5	p_6	p_{10}	p_{11}	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
4	1	8	6	7	0.25	0	0.75	0
8	3	3	1	2	0	1	0	0
4	2	8	7	6	0	0	1	0
3	7	2	8	7	0	0	0	1
		⋮	⋮	⋮	⋮	⋮	⋮	⋮

(b) Data using the time signal in level 1

p_1	p_5	p_8	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
8	2	5	0.15	0	0.85	0
8	6	1	0	1	0	0
8	2	6	0	0	1	0
5	8	8	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

(c) Data using the time signal in level 2

p_6	p_{10}	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
8	5	1	0	0	0
1	3	0	1	0	0
5	5	0	0	1	0
2	4	0	0	0	1
	⋮	⋮	⋮	⋮	⋮

(d) Data using the time signal in level 3

p_1	p_5	p_8	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
8	2	5	0.15	0	0.85	0
8	6	1	0	1	0	0
8	2	6	0	0	1	0
5	8	8	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

(e) Data using the time signal in level 4

p_3	p_4	p_{10}	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
6	6	6	0.15	0	0.85	0
8	1	3	0	1	0	0
1	2	2	0	0	1	0
4	5	4	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

(f) Data using the time signal in level 5

p_1	p_2	$^{N}\beta_{ri}$	$^{M}\beta_{ri}$	$^{U}\beta_{ri}$	$^{L}\beta_{ri}$
3	2	1	0	0	0
8	8	0	1	0	0
2	5	0	0	1	0
3	2	0	0	0	1
	⋮	⋮	⋮	⋮	⋮

(g) Data using the time signal in level 6

Figure. 10 Parts of the acquired knowledge of diagnosis for PNN learning

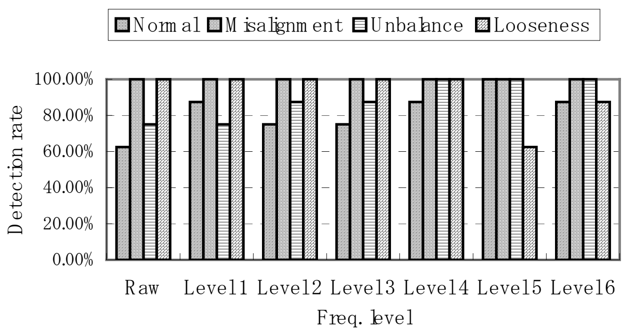

Figure.11 Comparison of detection rate in each state with each frequency level

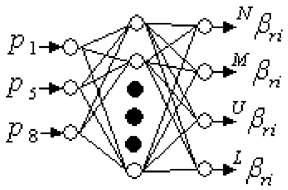

Figure. 12 The PNN for diagnosis using recomposed time signal in level 4

Fig. 12 shows the PNN built on the basis of the method proposed in this paper. By learning the knowledge shown in Fig. 10 (e), the PNN can quickly diagnose the faults with the possibility grade ${}^s\beta_{ri}$.

We sued the data measured in each state, which have not been learnt by the PNN, to verify the efficiency of the PNN. The data are time signals recomposed by wavelet transformation. When the data of symptom parameters measured in the normal state (N), are inputted to the PNN, the probability of correct judgment is 87.5%. Similarly, the probabilities of correct judgment for misalignment, unbalance and looseness stets are all 100.0% respectively as shown in Fig. 13.

State by judgment ↓	Sate of signals for verifications			
	N	M	U	L
Normal	87.5%	0.0%	12.5%	0.0%
Misalignment	0.0%	100.0%	0.0%	0.0%
Unbalance	0.00%	0.0%	100.0%	0.0%
Looseness	0.00%	0.0%	0.0%	100.0%

Figure. 13 Verification

7. Conclusions

In order to diagnosing faults of rotating machinery and processing ambiguous information for diagnosis by neural network, This paper proposed an intelligent diagnosis method for condition diagnosis of plant machinery using wavelet transform (WT), rough sets (RS) and partially-linearized neural network (PNN) to detect faults and distinguish fault type at an early stage. The WT is used to extract feature signal of each machine state from measured vibration signal for high-accurate condition diagnosis. The decision method of optimum frequency area for the extraction of feature signal was discussed using real plant data. We also proposed the diagnosis method by using "Partially-linearized Neural Network (PNN)" by which the fault types can be automatically distinguished on the basis of the probability distributions of symptom parameters. The symptom parameters are non-dimensional parameters which reflect the characteristics of time signal measured for condition diagnosis of plant machinery. The knowledge for the PNN learning can be acquired by using the Rough Sets (RS) of the symptom parameters. The diagnosis results of structural faults of rotating machinery have been shown to verify the effectiveness of the methods discussed here.

The superiority of the method proposed here can be explained as;

(1) Feature signal of each machine state can be extracted by the wavelet transformation for high-accurate condition diagnosis;

(2) The PNN can deal with the ambiguous problems of condition diagnosis;

(3) The PNN is always convergent because the learning data is made consistent by rough sets;

(3) The convergence of the PNN when learning is speedy.

References

[1] H. MATUYAMA: Diagnosis Algorithm, Journal of JSPE, Vol.75 No.3, pp.35-37, 1991.

[2] Christopher Bishop M., “Neural Networks for Pattern Recognition”, Oxford University Press, 1995.

[3] Daubechies, I., “Ten lectures on wavelets”, SIAM, pp.269-285, 1992.

[4] Pawlak Z., “Rough Sets”, International Journal of Computer Information Science. Vol.11, pp.344-356, 1982.