Journal of the Korean Nuclear Society
Volume 35, Number 5, pp.454~470, October, 2003

Fault Diagnosis for Agitator Driving System in a High Temperature Reduction Reactor

Gee Young Park, Dong Hee Hong, Jae Hoo Jung, Young Hwan Kim, Jae Hyun Jin, and Ji Sup Yoon
Korea Atomic Energy Research Institute
150 Dukjin-dong Yuseung-gu, Daejeon 305-353, Korea
gypark@kaeri.re.kr

(Received June 30, 2003)

Abstract

In this paper, a preliminary study for development of a fault diagnosis is presented for monitoring and diagnosing faults in the agitator driving system of a high temperature reduction reactor. In order to identify a fault occurrence and classify the fault cause, vibration signals measured by accelerometers on the outer shroud of the agitator driving system are firstly decomposed by wavelet transform (WT) and the features corresponding to each fault type are extracted. For the diagnosis, the fuzzy ARTMAP is employed and thereby, based on the features extracted from the WT, the robust fault classifier can be implemented with a very short training time - a single training epoch and a single learning iteration is sufficient for training the fault classifier. The test results demonstrate satisfactory classification for the faults pre-categorized from considerations of possible occurrence during experiments on a small-scale reduction reactor.

Key Words : agitator driving system, vibration signals, wavelet transform, feature extraction, fuzzy ARTMAP

1. Introduction

The objective of this research is the detection of an abnormal status of the agitator driving system in a high temperature reduction reactor in order to increase operational performance. Otherwise, the abnormal rotations may induce improper reduction reactions and also may interfere severely with the other monitoring system for reduction reactor integrity [1]. From this diagnosis result, the operator can recognize the integrity of the agitator driving system and might set up a schedule for repair or maintenance between operations.

In the high temperature reduction reactor, the metal reduction for nuclear spent oxide fuel takes place using a very reactive chemical de-oxidizer under a high temperature of 650℃. Fig.1 depicts the full-scale high temperature reduction reactor that was constructed last year but is not fully operational yet. The full-scale reduction reactor

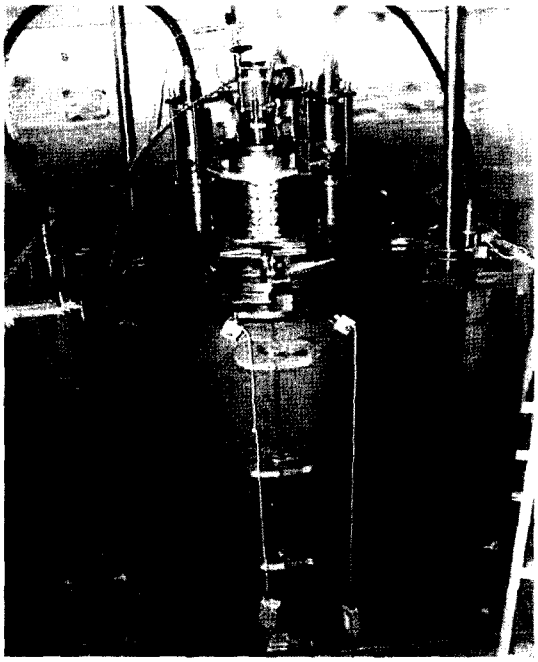

Fig. 1 Full-Scale Reduction Reactor

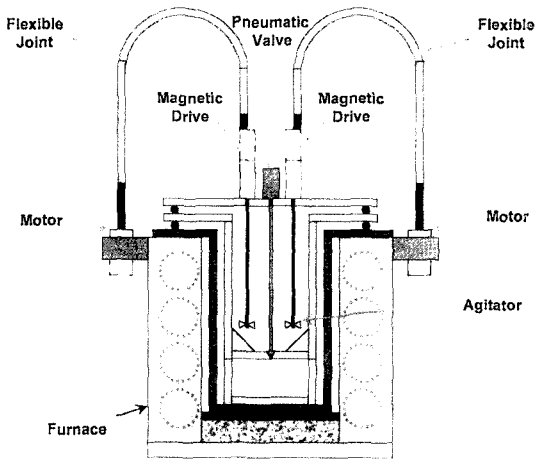

Fig. 2 Configuration of Reduction Reactor

has the dimension of 2.4 (H) × 1.7 (ID) m. Fig.2 shows the schematic of the configuration of the reduction reactor and the installation of the main components. The agitator driving system is composed of magnetic driver and agitator. In Fig.2, two magnetic drivers are installed on the top of the reduction reactor. Between those magnetic drivers is the needle valve operating equipment. At the temperature of 650℃, the reaction-supporting salt begins to melt down and, at that time, the agitator driving system is actuated for efficient chemical reactions. As the agitator driving system is installed on the top of the thermal reduction reactor, about half of the rotating axis and the agitator are immersed in the high-temperature molten salt. The magnetic driver is actuated by the motor, installed on the side of the reduction reactor as can be seen in Fig.2, whose rotating axis is arranged in a parallel direction to that of the magnetic driver. Both axes are inter-connected by flexible joint cable that bends in the upper half circle.

In the preliminary reduction experiments using uranium oxide powder, which commenced in 2001, on the small-scale reduction reactor that is an one-fifth scale of the full-scale reactor and on which the single agitator driving system with a smaller power capacity was installed, the power transmission via the flexible joint cable produced large vibrations in the structure. And, for a series of operations, this induced a loosening of the surrounding bolts. During the reduction operation at 650℃, the fumes generated within the reactor propagated through the internal space of the magnetic driver but the outer sealing mechanism of the magnetic driver prevented the fumes from leaking into the environment. After the shut down of the reduction reactor, the cooling down to the ambient temperature resulted in the accumulation of fumes in the clearance between the rotating axis and the outer ring, and this made the agitator driving system rotate intermittently with a sudden impact-like vibrations in the next operation. Moreover, the bearing in the outer case of the magnetic driver was sometimes exposed to the fumes emanating from the sample hole during

sampling of the salt, which resulted in the corrosion and wear of the bearings in the magnetic driver.

All the faults described above were witnessed independently in the preliminary experiments on the small-scale reduction reactor but not observed yet in the full-scale reduction reactor because only a few test operations have been performed while many experiments have been performed on the small-scale reduction reactor. These faults increase significantly the burden of regulating the rotating speed of the agitator to the pre-set point, 200 rpm, for the optimal reduction reaction and interfere severely with the other fault monitoring technique incorporated with the acoustic emission method [1][2]. In order to develop the diagnosis technique that recognizes fault occurrence and its cause, an experimental facility of the agitator driving system was constructed. This technique is based on the vibration signals measured by two accelerometers on the outer shroud of the magnetic driver.

2. Experimental Setup

For the acquisition of the vibration signals, a small-scale experimental facility for the agitator driving system was constructed. Fig.3 shows the experimental facility. The experimental setup is a mock-up of the small-scale reduction reactor in which the configuration of the agitator driving system is almost the same as that of the full-scale reactor.

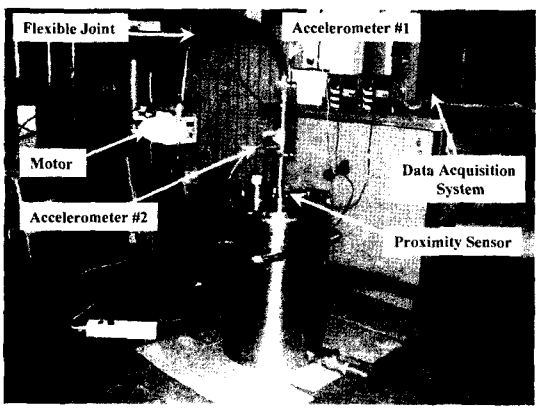

Fig. 3. Experimental Facility for Agitator Driving System

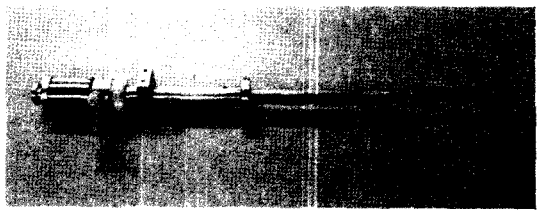

Fig. 4. (a) Magnetic Driver Manufactured

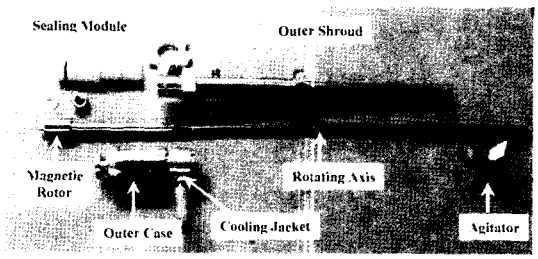

Fig. 4. (b) Components of Magnetic Driver

Fig.4(a) and Fig.4(b) show the magnetic driver manufactured and each component after it was dismantled, respectively. The magnetic driver in the agitator driving system has a magnetic pulley installed within the outer case and a magnetic rotor on the top of the rotating axis. The magnetic pulley rotates with the flexible joint cable and this rotation induces the actuation of the magnetic rotor by a magnetic force that allows the maximum torque of 6kgf·cm of the agitator attached at the bottom end of the rotating axis. The magnetic driver is installed on the top of the reduction reactor and has a supporting plate that supports the magnetic driver, tightening the joint part between the flexible cable end and the magnetic pulley. The supporting plate is connected to the outer shroud of the magnetic driver by two upper and lower bolts.

In the experiments, as the measuring elements, two accelerometers and two proximity sensors were installed on the agitator driving system. As can be seen in Fig.3, one accelerometer was attached to the top of the outer case of the magnetic driver and the other to the outer shroud enclosing the rotating axis. Two proximity sensors were installed on the outer shroud, penetrating through the outer shroud, about 2 mm away from the rotating axis within the outer shroud and arranged about 45° apart from each other. For proximity sensors, though installed and used in this experiment, we discarded the values of these sensors because there existed the possibility of the sealing problem due to the penetration of the proximity sensors through the outer shroud and the deposition of the fumes on each sensing end of the proximity sensors when they are implemented in the reduction reactor. The fact that none of the revolution measurements could be used made the data acquisition system never achieve the synchronous data acquisition where one data block exactly corresponds to single revolution of the rotating axis. This fact induced a severe restriction on the application of various methods to the feature extraction as mentioned in a later section. Moreover, between both accelerometers, the vibration signals from the accelerometer #2 as in Fig.3 was used in the signal analysis because it was found that this sensor was generally more sensitive to all the fault types.

In order to measure and analyze the vibration signals, a portable data acquisition system was constructed as shown in the upper-right part of Fig.3. The data acquisition system comprises of a digital signal processing board with various software-selectable sampling rates, signal conditioning modules, a main processing unit, and peripherals. The signal conditioning modules can accommodate 4 channels for the accelerometers and 2 channels for the proximity sensors. The signal conditioning modules for the accelerometers provide the anti-aliasing filter with 4 levels of cut-off frequency: 1, 5, 10, and 20 kHz. The accelerometers are charge-type and the effective frequency range is from 1 Hz to 10 kHz.

3. Data Acquisition for Fault Types

For the agitator driving system, the faults are classified grossly into three types: the looseness of tighteners, the bearing defects, and the clearance blocking. This classification is deduced from the observations of the occurrence of faults during the experiments with the small-scale reduction reactor. For the looseness, there are four bolts on the supporting plate and therefore, there are many cases for fault classification related to the looseness. In the experiments of the small-scale reduction reactor, the vibration forces by the non-rigid rotations of the flexible cable affected more severely the upper bolts than the lower bolts and hence, the looseness was mainly observed at the upper bolts. Thus, for the simplicity in presenting the results, the looseness of the two upper bolts is only considered. The looseness features of both the left and right bolts are same and these effects are combined into a single category. The fault types categorized are presented in Table 1.

In the conventional monitoring/diagnosis techniques for the rotary machines, the

Table 1. Fault Types Categorized

Fault Type	Contents
Case 1	Normal Rotation
Case 2	Looseness of One of Upper Bolts
Case 3	Looseness of Upper Bolts
Case 4	Bearing Defects
Case 5	Clearance Blocking

synchronization between the revolution of rotating axis and the data acquisition is very important for delicate analysis and identification of an incipient fault. For the agitator driving system in the high temperature reduction reactor, there is no way to achieve synchronization because motors and magnetic drivers constructed in the high temperature reduction process are not provided with any revolution measuring unit and it is not reasonable to install additional rotation-measuring sensors because of the sealing problem. Moreover, the motor has no feedback controller that regulates strictly the rotation speed to the predesignated value of 200 rpm and hence, the rotation speed is actually fluctuating around 200 rpm during the operation, which prohibits the setting of a proper sampling rate of the data acquisition system for the synchronization. As a result, it is impossible to synchronize the data acquisition and the revolution of the rotating axis in this facility.

The sampling rate is set to 25.6 kHz, which is very much higher than the maximum frequency of the agitator axis rotations but is necessary to include the frequency components of the transient signals by a fault or the flexible joint cable. The rotation speed is fluctuated with a maximum deviation ± 20 rpm from the set-point speed. The number of samples in the data block is set by 16384 enough for encompassing sufficiently the signals per revolution under the speed fluctuations from 180 to 220 rpm. Through the experiments, the vibration signals for various set-point speeds from 80 to 800 rpm and for various faults were measured and analyzed. In this paper, among all experimental data, the data measured at the set-point speed of 200 rpm are used in the diagnosis procedure because this rotation speed of 200 rpm is a set point value employed in reduction reactor operations.

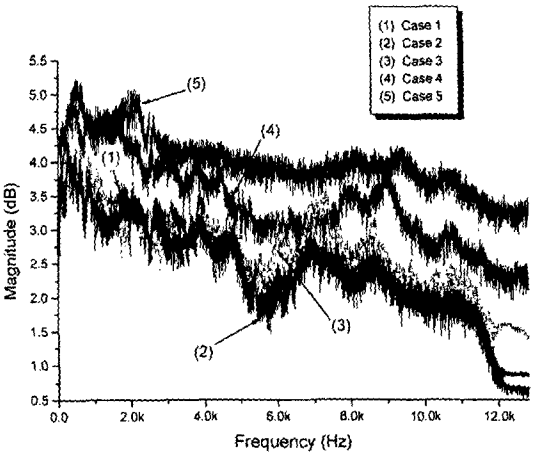

Fig. 5. Averaged Power Spectrum of FFT for 5 Types of Faults

4. Configuration of Diagnosis System

With the acquired data blocks, the FFT (Fast Fourier Transform) method is applied to the analysis of signals for pre-defined fault types. Fig.5 shows the power spectrum of 5 different cases by the FFT with 8 averages of FFT results for eliminating highly fluctuating components in each spectrum. As can be seen in Fig.5, some faults such as the clearance blocking and the bearing defects show slightly different trends from the normal rotation but it is not easy to identify the distinguishable peaks representing the corresponding fault.

In order to overcome the limitation of the conventional signal analysis method, a fault diagnosis system for the agitator driving system is developed and its configuration is depicted in Fig.6. The data blocks of vibration signals are stored in the data acquisition system and, after the data storage is completed, these data blocks are transferred to the wavelet transform (WT) as input data. The WT decomposes these signals into the time- and scale-domain signals and the features are extracted from the signals across the scale.

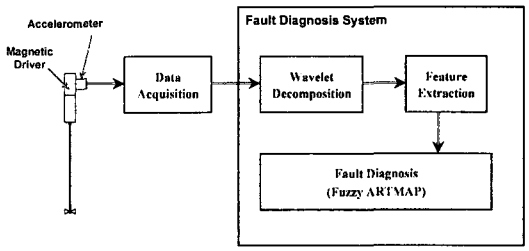

Fig.6 Configuration of Fault Diagnosis System

The fault diagnosis is performed by use of the fuzzy ARTMAP based on the features extracted. The fuzzy ARTMAP should be trained before performing fault diagnosis or classification. The learning speed and the training time of the fuzzy ARTMAP are much faster than that of any other artificial neural network architecture for classification.

5. Brief Description on the Wavelet Transform

Let $L_2(R)$ denote the Hilbert space that is the set of square integrable functions. Then, the continuous wavelet transform of a function $f \in L_2(R)$ is defined by [3][4]

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)^* dt. \quad (1)$$

In Eq.(1), W represents the wavelet transformation and the superscript * means the complex conjugation. $\psi_{a,b}$ is the wavelet represented as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \text{ for } (a, b) \in R^2 \text{ and } a>0, \quad (2)$$

where, $\psi(t)$ is a fixed function called the mother wavelet. The parameter a is a scaling factor and it stretches or compresses the mother wavelet ψ. The parameter b is a translation along the time axis. It simply shifts a wavelet, i.e., it delays and advances the time at which it is activated. The factor $\sqrt{a}$ ensures that the functions $\psi_{a,b}$ have a constant norm in the space $L_2(R)$. The scaling operation on the mother wavelet enables the wavelets to capture the different frequency components of the function to be analyzed. The translation operations shift the mother wavelet in the time axis in order to capture local time information of the function. The wavelets produced by Eq.(2) thus provide the analysis of the function at different times and at different scales (or, roughly, frequencies).

The Fourier transform (FT) is the most frequently used method in signal processing and analysis. If we are given a single frequency signal $e^{i\omega t}$, the FT will isolate a peak at the frequency ω. However, when confronted with the case of a signal built of two single frequencies occurring at two adjacent intervals, we obtain two peaks without localization of time. The resolution in the frequency domain can shrink to one isolation point but the time domain resolution is the entire range. One of the tools for localization both in time and frequency is the short-time FT. In this representation, one has to fix both the localization levels of time and frequency. Wavelets offer a different compromise: the frequency localization is logarithmic, i.e., proportional to the frequency level and the time localization gets finer in higher frequencies.

From the WT, the function $f(t)$ can be reconstructed by

$$f(t) = C_\psi^{-1} \int_0^\infty \int_{-\infty}^{\infty} W(a,b)\psi_{a,b}(t) db \frac{da}{a^2}. \quad (3)$$

And, in order for Eq.(3) to be valid, the wavelet must satisfy the following admissibility condition [3]

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} \le \infty, \quad (4)$$

where Ψ is the FT of the wavelet ψ. Note that the

condition Eq.(4) implies that $\Psi(0) = 0$ so that ψ is a basically oscillating function, i.e., $\int\psi dt = 0$. More generally, one can design ψ to have

$$\left.\frac{d^k\Psi(\omega)}{d\omega^k}\right|_{\omega=0} = 0 = \int t^k\psi(t)dt, \text{ for } k=1,\cdots,N. \quad (5)$$

An important consequence of making ψ have the property of Eq.(5) is that the values of W(a,b) are almost influenced by the regularity of the function. In the continuous WT, the parameters (a, b) are varied continuously. For implementation, the continuous WT should be computed on a discrete grid of points such as $(a_0^n, ma_0^n b_0)$ with $(n, m) \in Z^2$, and this makes Eq.(2) become

$$\psi_{n,m}(t) = a_0^{-n/2}\psi\left(a_0^{-n}t - b_0 m\right), \text{ for } a_0 > 0,\ b_0 \neq 0.$$

This means the sampling of the continuous WT is represented by

$$W(n,m) = \int_{-\infty}^{\infty} f(t)\psi_{n,m}(t)^* dt.$$

In order to preserve the information of $f(t)$ in the samples of the continuous WT, one should over-sample the continuous WT on a very dense grid, which requires considerable computational efforts. By adding restrictions on the parameter a_0 such that this parameter should have values determined by the powers of 2, i.e., dyadic values and also by the choice of the orthonormal wavelet, it is possible to remove redundancy and to construct a fast and efficient algorithm. In this case, the wavelets are represented as, by setting $a_0=2$ and $b_0=1$,

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right), \text{ for } (j,k)\in Z^2. \quad (6)$$

The WT by use of Eq.(6) is called the discrete wavelet transform (DWT). For discrete time sequences, the WT decomposes the sequences by using the octave-band filter banks when the dyadic and orthonormal wavelets are employed. The fast and efficient algorithms for the WT of discrete samples have been proposed [5][6].

5.1. Signal Decomposition by Wavelet Transform

For the agitator driving system to be monitored and diagnosed, the vibration signals are decomposed into the octave band components by the DWT. The wavelet used in this decomposition is the so-called 'db-10' wavelet proposed by Daubechies [3]. In the db-10 wavelet, both low-pass and high-pass FIR filters have 20 filter coefficients. This requires more calculations but the sharp attenuation at the cut-off frequency of each octave-band filter can be achieved.

Fig.7(a) through Fig.7(e) show the results of the wavelet decomposition of the five types of faults. The data block is decomposed into 8 scale levels, which results in 8 details and 1 approximation. The term 'detail' means, roughly speaking, that a data block with a certain frequency range is decomposed into the signals whose frequency components are above half of the frequency range of a data block. The decomposed signals below the half frequency range are called the 'approximation'. In Figs.(7), the original data is displayed in the top-left graph and the other graphs are the decomposed signals. The signal with the highest frequency range is depicted in the bottom-right graph as indicated by 'Detail 1'. As the number of 'Detail' increases, the signal contains lower and lower frequency components and, in the top-right graph marked as 'Approximation', the lowest frequency components are contained in the signal.

As can be seen in Fig.7(a) to Fig.7(e), the decomposed signals show different characteristics for corresponding faults but it is very difficult to evaluate those differences compactly and

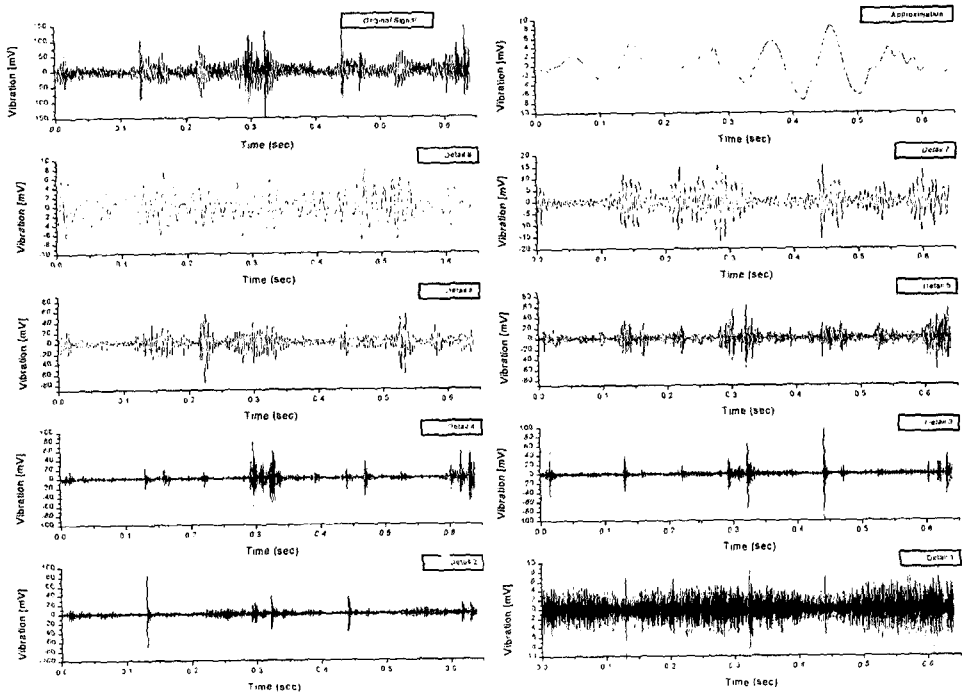

Fig. 7(a). Wavelet Transform for Case 1

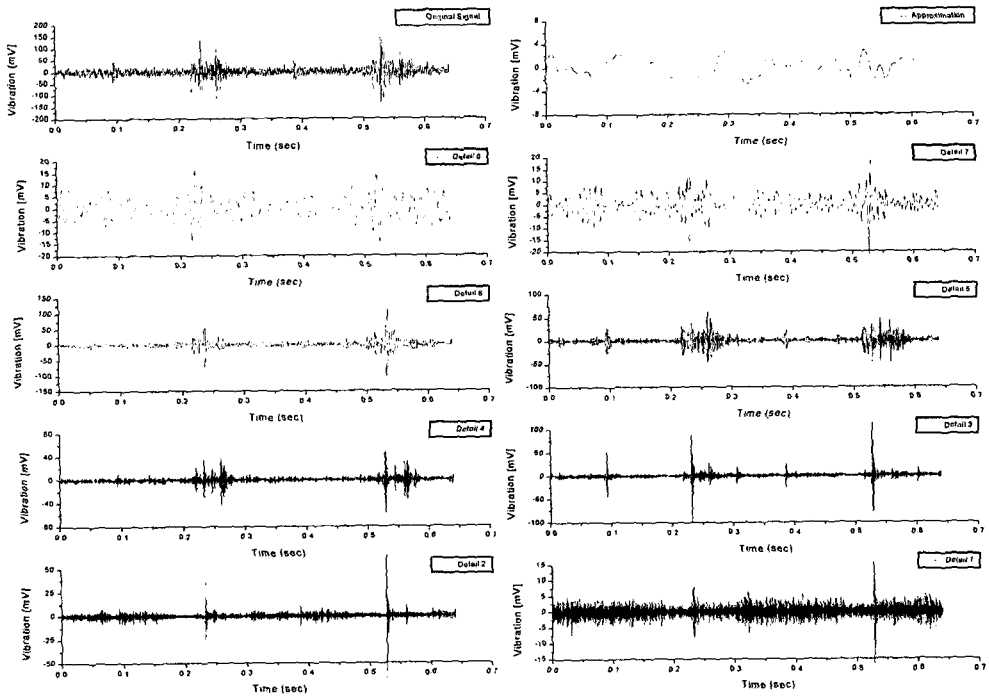

Fig. 7(b). Wavelet Transform for Case 2

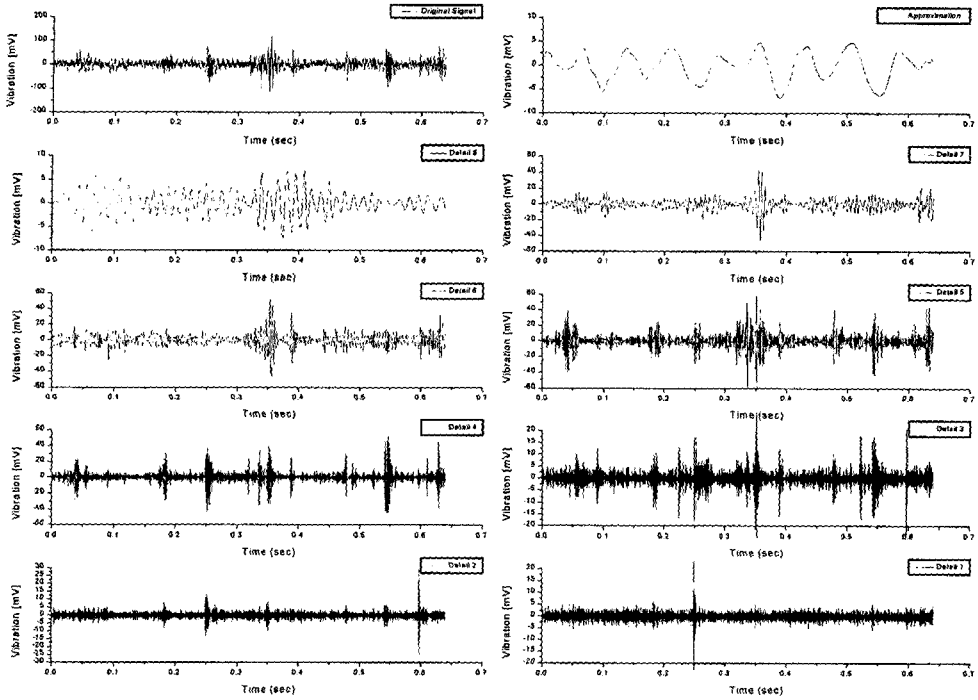

Fig. 7(c). Wavelet Transform for Case 3

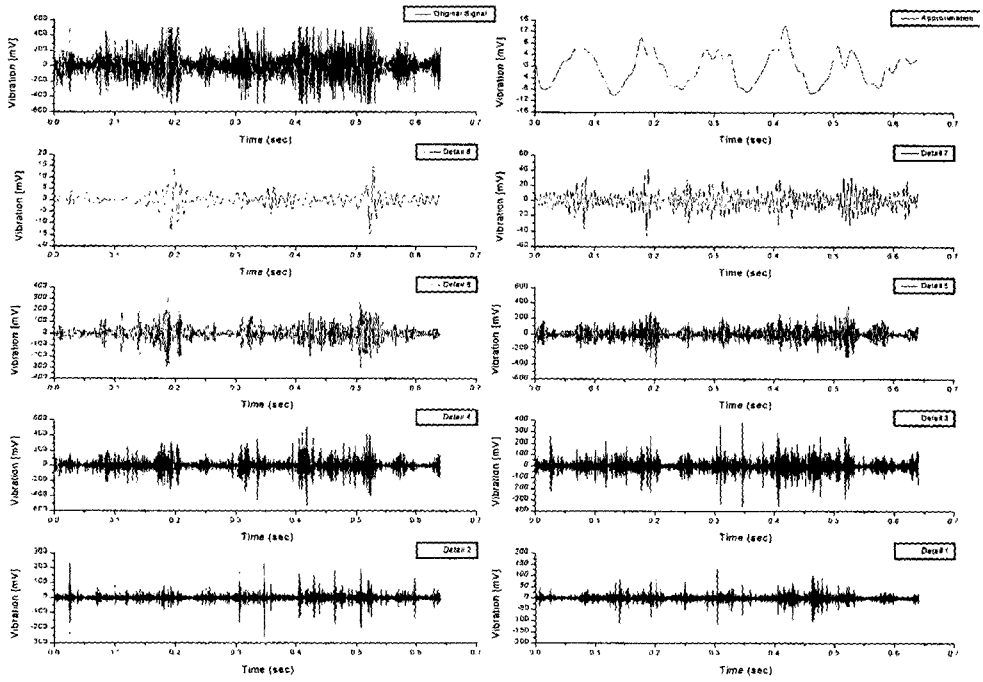

Fig. 7(d). Wavelet Transform for Case 4

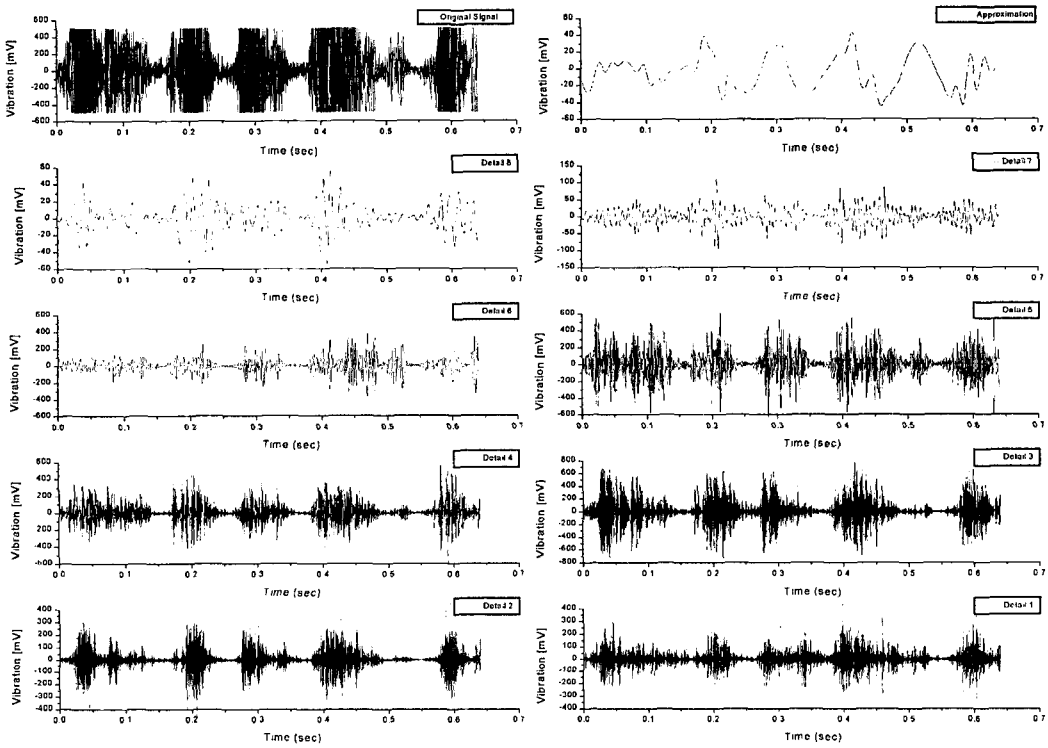

Fig.7(e). Wavelet Transform for Case 5

systematically from the decomposed signals by the WT. It is desired, therefore, that the features in the decomposed signals for a type of fault be extracted in such a way that the features are compact enough and represent efficiently the information of the fault type.

6. Feature Extraction

Based on the results of the wavelet decomposition for the vibration signals, the appropriate and compressed features are extracted. The feature extraction has been performed by use of various methods such as AR modeling, principal component analysis by an artificial neural network, etc. All those methods have not produced a good and consistent feature extraction for a type of fault because the data blocks for any type of fault are not synchronized.

Thus, in this paper, a simple method is employed for feature extraction. The feature values for eight details and one approximation for a type of fault are obtained by calculating the average power for each detail and approximation signals, and this is represented by

$$F_j = \sum_{i=0}^{N-1} | d_{j,i} |^2 / N, \text{ for } j = 1, 2, \cdots, J, \text{ and}$$

$$F_{J+1} = \sum_{i=0}^{N-1} | a_{J,i} |^2 / N,$$

where F_j represents the feature value at the scale level of j, $d_{j,i}$ is the value of the wavelet transform or decomposed signal at the scale level of j and at the position i, $a_{J,i}$ is the value of the approximation at the scale level of J and at the position i, N

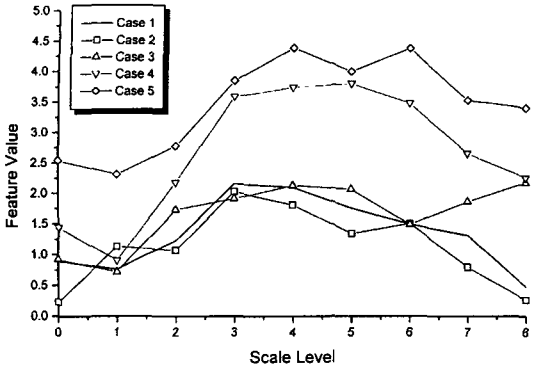

Fig.8 Result of Feature Extraction

represents the total number of a data block (N = 16384), and J is the maximum scale level (J = 8). For the display on the graph, the feature vector $\mathbf{V}_F$ is defined by

$$\mathbf{V}_F = [V_{F,1}, \cdots, V_{F,J+1}] = [\log(F_{J+1}), \log(F_J), \log(F_{J-1}), \cdots, \log(F_1)]. \quad (7)$$

In the displays shown below, the feature values are plotted in the order of the components array as in Eq.(7). Fig.8 shows the results of the feature extraction by the average power calculation. As can be seen in Fig.8, a fault can be discriminated from the others when one inspects carefully the trend and characteristics of the feature values. In order to demonstrate the consistent trend of the feature vectors for a fault type, Fig.9(a) through Fig.9(e) show the features extracted from some, further acquired, data blocks. As can be seen in Figs.9, the features of further data blocks for a corresponding fault show a consistent trend that can be differentiated from the feature trends of the other faults. As a matter of fact, this method can also be applied to the results of FFT but the features from the wavelet transform are more distinguishable than that from FFT.

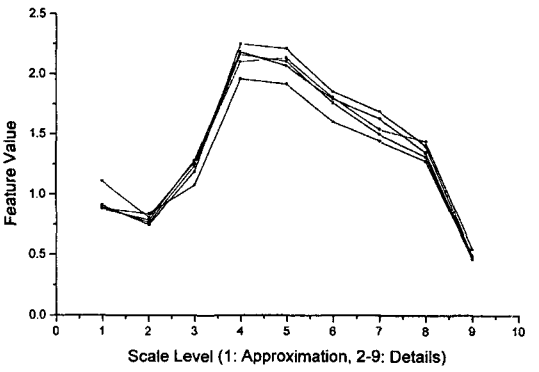

Fig. 9(a). Features from WT for Case1 Data Blocks for Test

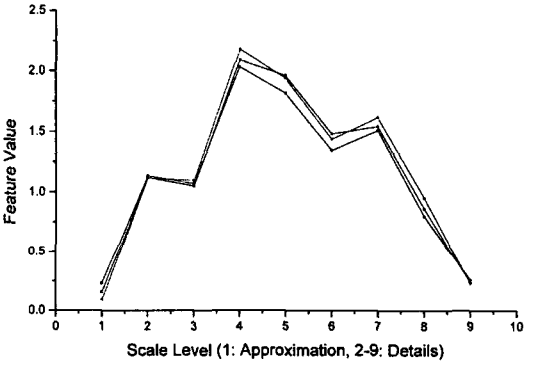

Fig. 9(b). Features from WT for Case 2 Data Blocks for Test

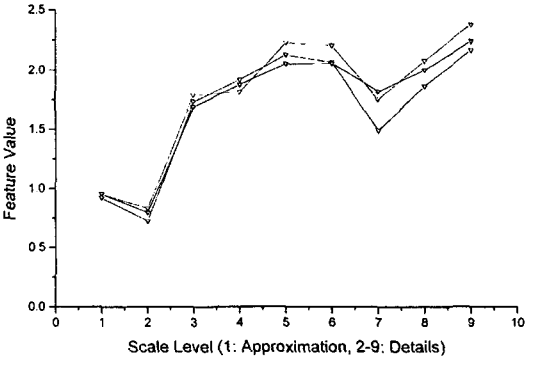

Fig. 9(c). Features from WT for Case 3 Data Blocks for Test

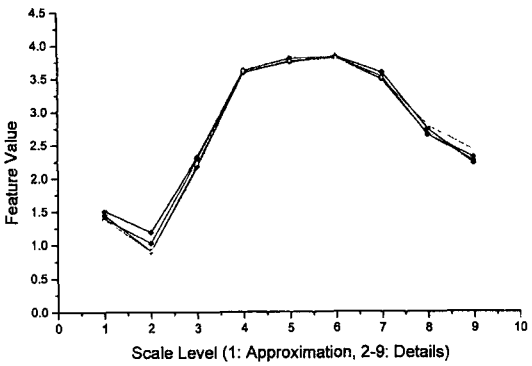

Fig. 9(d). Features from WT for Case 4 Data Blocks for Test

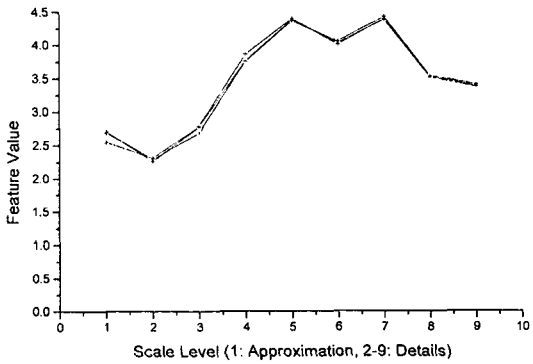

Fig. 9(e). Features from WT for Case 5 Data Blocks for Test

7. Diagnosis

For the diagnosis based on the features extracted from the average power calculation, the neural network classifier based on the adaptive resonance theory (ART) [7], which is called the fuzzy ARTMAP [8] in neural network literatures, is constructed.

7-1. Description of Fuzzy ARTMAP

The ART has the architecture where a new feature vector is learned without modifying the existing information for different features. Hence, it can preserve the previously learned knowledge while continuing to learn new things. The fuzzy ARTMAP is a class of artificial neural networks that perform incremental supervised learning of recognition categories and multidimensional maps in response to input vectors presented in arbitrary order. Fig.10 shows the architecture of the fuzzy ARTMAP where two fuzzy ART modules (ART_a and ART_b) and a map field F^{ab} are involved.

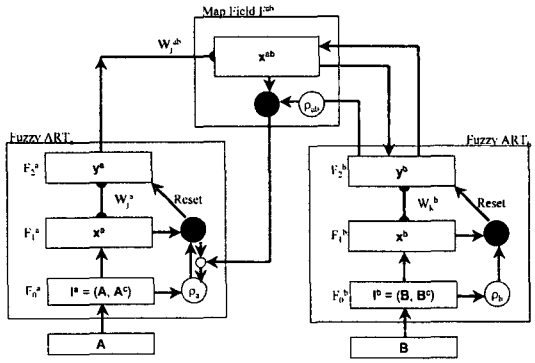

Fig. 10. Configuration of Fuzzy ARTMAP

The input vectors **A** of dimension M_a and **B** of dimension M_b are the feature vectors respectively corresponding to the symptom and the cause. The components of each input vector are analog or binary values within the range of [0,1]. Each component of the input vector represents a feature value or a feature item that is grouped to establish the representative feature vector. Thus the analog value of each component of the input vector means the degree of belongingness to the corresponding feature item, and this is similar to the fuzzy membership value. Given the input vectors **A** and **B**, the overall operations of the fuzzy ARTMAP are described below.

When input vectors are given, the input vector **A** (or **B**) is transferred to the F_0^a (F_0^b) layer where the complement coding is performed based on the input vector. In the F_0^a layer of Fig.10, $\mathbf{A}^c$ means

the complement of **A**, that is, at ith component of $\mathbf{A}^c$, $A_i^c = 1 - A_i$ for $i = 1, \cdots, M_a$. By the complement coding, the dimension of the input vector $\mathbf{I}^a$ increases by $2M_a$ and the magnitude of the complement-coded input vector $\mathbf{I}^a$ is always constant to M_a regardless of the input vector **A**. This complement coding is important in the fuzzy ART architecture because it prevents the so-called category proliferation due to the decrease of the weight vector $\mathbf{W}_j^a$ [9]. The $\mathbf{W}_j^a$ is an $1 \times 2M_a$ vector such that $W_j^a = [W_{j,1}^a, \cdots, W_{j,2Ma}^a]$ and the subscript j represents *j*th node at the F_2^a layer. All initial values of $\mathbf{W}_j^a$ are set to 1. After complement coding, $\mathbf{I}^a$ is presented to the input terminal of the F_1^a layer and this becomes $\mathbf{x}^a$. This interim product $\mathbf{x}^a$ is fed up to the F_2^a layer through the weight vector $\mathbf{W}_j^a$ for $j = 1, \cdots, N_a$, where N_a is the total number of the nodes in the F_2^a layer.

For the input vector $\mathbf{I}^a$ and a node j at the F_2^a layer, the choice function, T_j, is defined by

$$T_j(\mathbf{I}^a) = |\mathbf{I}^a \wedge \mathbf{W}_j^a| / (\alpha + |\mathbf{W}_j^a|), \text{ for } j = 1, \cdots, N_a. \quad (8)$$

In Eq.(8), α is called the conservative parameter that has a very small value. And, for M dimensional vectors **p** and **q**, $\mathbf{p} \wedge \mathbf{q} = \{\min(p_i, q_i)$: for $\forall i\}$ and $|\mathbf{p}| = \Sigma_{i=1}^{M} |p_i|$. Among the values of T_j ($j = 1, \cdots, N_a$), *the winning node indexed by J is selected* by $T_J = \max\{T_j: j = 1, \cdots, N_a\}$. The output vector $\mathbf{y}^a$ is obtained by setting all the components of $\mathbf{y}^a$ to zero but Jth component given to one. Then, this output vector y^a goes down to the F_1^a layer where $\mathbf{x}^a$ now becomes $\mathbf{x}^a = \mathbf{I}^a \wedge \mathbf{W}_J^a$.

If the following condition is satisfied,

$$|\mathbf{x}^a| = |\mathbf{I}^a \wedge \mathbf{W}_J^a| \geq \rho_a |\mathbf{I}^a|, \quad (9)$$

then it is said the resonance state occurs and the learning is performed by

$$\mathbf{W}_J^a(\text{new}) = \beta(\mathbf{I}^a \wedge \mathbf{W}_J^a(\text{old})) + (1-\beta)\,\mathbf{W}_J^a(\text{old}). \quad (10)$$

In Eq.(10), β is a constant in the range of [0,1] and ρ_a in Eq.(9) is called the vigilance parameter having a value of 0~1. Initially, ρ_a is set to an arbitrary low value, ρ_a^{base}. If the condition in Eq.(9) is not satisfied, the reset operation occurs in such a way that the value of y_J^a is set to zero and *J*th node of the F_2^a layer is prohibited from being the winning node for further presentation of the input vector $\mathbf{I}^a$. And the input vector $\mathbf{I}^a$ is presented again and the calculations described above are iterated until the resonance state occurs. The operations of the fuzzy ART_b, when the input vector **B** is given and the number of the output node is set to N_b, are the same as in the fuzzy ART_a. The wining node at the F_2^b layer is indexed by K.

The map field F^{ab} is activated whenever one of the fuzzy ART modules (fuzzy ART_a and ART_b) is in the resonance state. If a node J of the F_2^a layer is chosen at the resonance state of the fuzzy ART_a, then weight $\mathbf{W}_J^{ab}$ ($\mathbf{W}_J^{ab} = [W_{J,1}^{ab}, \cdots, W_{J,Nb}^{ab}]$) activates the map field F^{ab}. If the fuzzy ART_b is active, i.e., in the resonance state, at a node K of the F_2^b layer, then the node K in the F^{ab} is activated by one-to-one pathways between F_2^b and F^{ab}. If both fuzzy ART_a and ART_b are active, then the F^{ab} becomes active only if the fuzzy ART_a predicts the same category as the fuzzy ART_b via the weight vector W_J^{ab} (that is, $W_{J,K}^{ab} \neq 0$). The output vector $\mathbf{x}^{ab}$ at the map field F^{ab} is calculated by

$$\mathbf{x}^{ab} = \begin{cases} \mathbf{y}^b \wedge \mathbf{W}_J^{ab} & F_2^a \text{active}, F_2^b \text{active} \\ \mathbf{W}_J^{ab} & F_2^a \text{active}, F_2^b \text{inactive} \\ \mathbf{y}^b & F_2^a \text{inactive}, F_2^b \text{active} \\ \mathbf{0} & F_2^a \text{inactive}, F_2^b \text{inactive} \end{cases}$$

With the values of $\mathbf{x}^{ab}$ for both layers F_2^a and F_2^b active, if the following condition is satisfied,

$$|\mathbf{x}^{ab}| \geq \rho_{ab} |\mathbf{y}^b|, \quad (11)$$

then, the mapping or prediction procedure is

completed. If the condition Eq.(11) is not satisfied, the vigilance parameter ρ_a in the fuzzy ART_a is increased from ρ_a^{base} by a small value in order for $|\mathbf{I}_a \wedge \mathbf{W}_J^a| < \rho_a |\mathbf{I}^a|$ to occur, which ultimately leads to activation of another F_2^a node J satisfying Eq.(9) and Eq.(11) for the input vector $\mathbf{I}^a$.

Initially, all components in $\mathbf{W}_j^{ab}$ (for $j = 1, \cdots, N_a$) are set to 1. When F_2^a node J and F_2^b node K are active under the training phase, then $W_{J,K}^{ab} = 1$ and $\mathbf{W}_{J,k}^{ab} = 0$ (for $k = 1, \cdots, N_b$ and $k \neq K$). When the training for particular input vectors **A** and **B** is completed, the fuzzy ARTMAP configures the weight vectors, $\mathbf{W}_j^a$, $\mathbf{W}_k^b$, and $\mathbf{W}_j^{ab}$. For another input presentation, the fuzzy ARTMAP assigns other winning nodes J and K at F_2^a and F_2^b layers, respectively, and stores the characteristics of both input vectors on the weight vectors $\mathbf{W}_j^a$ and $\mathbf{W}_k^b$, and matches the information from both the F_2^a and F_2^b layers by adjusting the weight vector $\mathbf{W}_j^{ab}$ at the map field F^{ab}.

After the training phase is finished, the classification test may start. At the test phase, the only input vector A′ is given to the F_0^a layer and then, the fuzzy ARTMAP goes to the fuzzy ART_b, through the Jth winning node in the F_2^a layer and the map field. The resonance state in the fuzzy ART_b always occurs in the test phase whenever some triggering value comes from the map field. Through the already configured $\mathbf{W}_k^b$, the triggering value from the map field goes down to the F_0^b layer where some nodes become the active state having a non-zero value. Eliminating the complement-coded nodes in the F_0^b layer results in the activated node(s) representing obviously the fault type in our case.

7.2. Diagnosis Test

For training and testing the fuzzy ARTMAP for diagnosing the agitator driving system, the five types of data sets are acquired. These data sets are measured when the corresponding fault is introduced intentionally in the experimental facility. For off-line training of the fuzzy ARTMAP, one data block for each fault is selected and the total training data are 5 sets as presented in Table 2. Fig.8 also shows the feature vectors of the selected training data set. The remaining data sets are used to test the fuzzy ARTMAP diagnosis performance. Total test data are 22 sets and the number of data blocks for each fault type is given in Table 2. Fig.9(a) to Fig.9(e) show the feature vectors of all the fault types used in the test phase. In the training phase, the training data pairs for each fault are presented just once to both input ports of the fuzzy ARTMAP and the single learning iteration is selected. The parameters in the fuzzy ARTMAP are given in Table 2 with short descriptions of the parameters.

Since the input vector A for the fuzzy ART_a is the feature values, the input vector **A** has a dimension of $M_a = 9$. All input vectors for the

Table 2. Data Sets and Configuration of Fuzzy ARTMAP

Training Data Sets	♦ Data Set : 5 Sets - Case 1 (Normal Rotation): 1 Set - Case 2 (One Upper Bolt Looseness): 1 Set - Case 3 (Upper Bolts Looseness): 1 Set - Case 4 (Bearing Defect): 1 Set - Case 5 (Clearance Blocking): 1 Set
Test Data Sets	♦ Data Set : 22 Sets - Case 1 (Normal Rotation): 7 Sets - Case 2 (One Upper Bolt Looseness):4 Sets - Case 3 (Upper Bolts Looseness): 3 Sets - Case 4 (Bearing Defect): 4 Sets - Case 5 (Clearance Blocking): 4 Sets
Training Type & Parameters	• Off-Line Learning, Single Input Presentation, and Single Learning Iteration • Fast Learn: $\beta = 1$ • Conservative Limit Value: $\alpha = 0.0001$ • Vigilance and Matching Criterion: $\rho_a = 0.8$, $\rho_b = 0.8$, and $\rho_{ab} = 0.8$

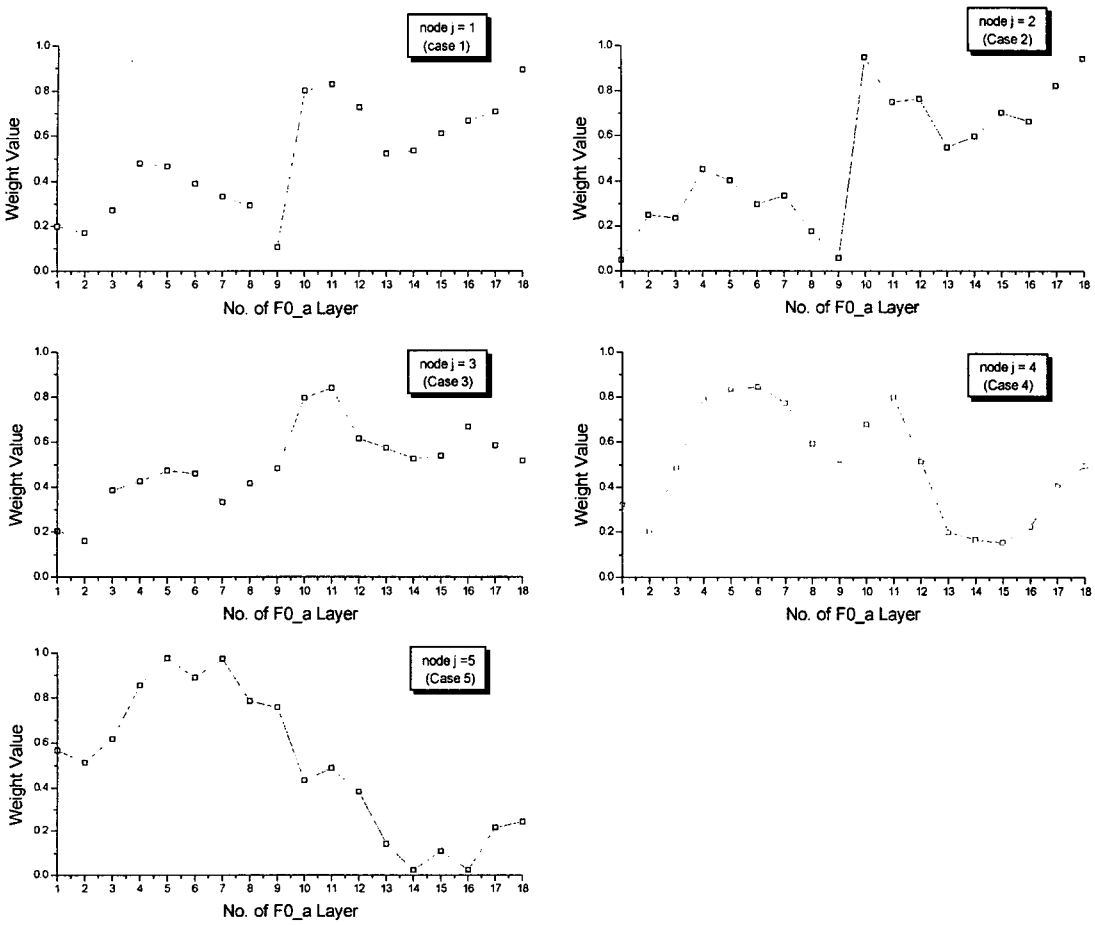

Fig.11. Information on Wja of Fuzzy ARTMAP

fuzzy ART_a are normalized into [0,1] by a constant value of 4.7. The input vector **B** for the fuzzy ART_b contains the binary values with a single '1' that represents a specific fault occurrence and has a dimension of $M_b = 6$.

After the training, the information established on the weight vector $\mathbf{W}_j^a$ is shown in Fig.11. Table 3 shows the constructed weight vector $\mathbf{W}_k^b$ of the fuzzy ART_b where the row indexes the nodes in the F_0^b layer and the column represents only the activated or winning nodes in the F_2^b layer. The values in W_k^b for a node k, actually, are the same as the input values in $\mathbf{I}^b$ that represent a fault type, i.e., at the node k = 1, $\mathbf{W}_1^b$ = [1,0,0,0,0,0,1,1,1,1] represents the fault type of Case 1. The weight vector $\mathbf{W}_j^{ab}$ becomes, in this case, a 5×5 identity matrix with j-indexed rows and k-indexed columns.

The learning time for the fuzzy ARTMAP at the training phase is extremely short because of the single learning iteration (β=1) and the single training epoch. After the training phase, the fuzzy ARTMAP shows the perfect classification for the training data set when the same input vectors are presented at the test mode.

For the test data sets, the fuzzy ARTMAP shows the correct fault identification performance for all

Table 3. The Information of W_k^b

Input Node / Node k	1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	0	0	1	1	1	1
2	0	1	0	0	0	1	0	1	1	1
3	0	0	1	0	0	1	1	0	1	1
4	0	0	0	1	0	1	1	1	0	1
5	0	0	0	0	1	1	1	1	1	0

the fault types though some test data have a little distorted trend as can be seen in Figs.9.

8. Conclusions and Further Research

In this paper, a fault monitoring technique is presented in order for monitoring and diagnosing the status of the agitator driving system in the high temperature reduction reactor. Due to the high-temperature chemical reduction reaction, the agitator driving system has the faults of clearance blocking and bearing defects by the fumes. And it has the faults of the loosening of some bolts by the non-rigid rotation transmission of the flexible joint cable. The faults in the agitator driving system affect the chemical reduction reaction and also interfere severely with the other fault monitoring technique incorporated with the acoustic emission method. In order to recognize the fault occurrence and the cause of that fault, a fault monitoring technique for the agitator driving system was developed. In order to identify the occurrence of a specific fault, the vibration signals for the corresponding fault are acquired and analyzed by wavelet decomposition. To extract the feature from the result of the wavelet decomposition, the average power calculation for each decomposed signal was employed.

For the diagnosis procedure based on the features, the neural network classifier based on the adaptive resonance theory, which is the Fuzzy ARTMAP, was constructed. The tests on the various fault signals including the normal operating signal show the perfect classification performance of the Fuzzy ARTMAP though the features of the experimental data from the same fault are varied or distorted.

Currently, the technique developed in this paper is being implemented into the data acquisition system for on line diagnosis and will be applied to the full-scale reduction reactor. In the real implementation, all cases for the looseness of the tighteners are to be combined into a single case, just differentiating this case from the other cases of the normal rotation, bearing defects, and clearance blocking. And all relevant works for acquiring the information of all the looseness cases are now being performed. As a further study, the analysis of multiple faults occurrence is to be considered when combined faults are generated in the further operations of the full-scale reduction reactor, though not observed up to this time. A delicate method for identifying the degradation level of a fault is supposed to be required and will be pursued in the further study.

References

1. G. Y. Park, et al., *Development of Fault Signal Analysis Technique for Advanced Spent Fuel Conditioning Process*, KAERI/TR-2156/2002, Technical Report, Korea Atomic Energy Research Institute, March (2002). (in Korean)
2. G. Y. Park, B. S. Park, J. S. Yoon, and D. H. Hong, "Development of Fault Monitoring System for Advanced Fuel Conditioning Process", *PBNC 2002 (The 13th Pacific Basin Nuclear Conference)*, Session:TS-9B, no.3, pp.1-8, Oct. 21~25, Shenzhen, China (2002).
3. I. Daubechies, *Ten Lectures on Wavelets*,

SIAM, Philadelphia, PA (1992).
4. C. S. Burrus, R. A. Gopinath, and H. Guo, Introduction to Wavelets and Wavelet Transforms, Prentice Hall (1998).
5. S. G. Mallat, "Multi-Frequency Channel Decompositions of Images and Wavelet Models", *IEEE Trans. Acoust., Speech, Signal Processing*, 37, pp.2091-2110 (1989).
6. M. J. Shensa, " The Discrete Wavelet Transform: Wedding the A Trous and Mallat Algorithms" , *IEEE Trans. on Signal Processing*, 40, no.10, pp.2464-2482 (1992).
7. G. A. Carpenter and S. Grossberg, "The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network", *Computer*, pp.77-88, March (1988).
8. G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multi-Dimensional Maps", *IEEE Trans. on Neural Networks*, 3, no.5, pp.698-713 (1992).
9. G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System" , *Neural Networks*, 4, pp.759-771 (1991).