



# EMPIRICAL MODE DECOMPOSITION AND ROUGH SET ATTRIBUTE REDUCTION FOR ULTRASONIC FLAW SIGNAL CLASSIFICATION

Yu Wang

School of Software Engineering  
Chongqing University of Arts and Sciences  
Chongqing 402160, China  
Email: llylab@126.com

---

*Submitted: Apr. 15, 2014*

*Accepted: July 30, 2014*

*Published: Sep. 1, 2014*

---

*Abstract- Feature extraction and selection are the most important techniques for ultrasonic flaw signal classification. In this study, empirical mode decomposition (EMD) is first used to obtain the intrinsic mode functions (IMFs) of original ultrasonic signals. Such IMFs and traditional time as well as frequency domain based statistical parameters are extracted as the initial features of flaw signal. After that, spectral clustering method is used for feature value discretization so that rough set attribute reduction (RSAR) can be applied to implement feature selection. Finally, the selected features are taken as input of artificial neural networks (ANNs) to train the decision classifier for flaw identification. Experimental results show that compared to conventional wavelet transform based schemes and principal components analysis, EMD combined with RSAR can improve the performance of feature extraction and selection. Using such hybrid scheme can effectively classify different ultrasonic flaw signals with high accuracy and low training elapsed time.*

**Index terms:** Empirical mode decomposition, rough set attribute reduction, feature extraction and selection, ultrasonic flaw signal classification

## I. INTRODUCTION

Ultrasonic inspection is one of the most successful non-destructive testing (NDT) techniques for quality assessment and detection of flaws in engineering materials since it can analyze even highly thick metallic as well as nonmetallic materials with good resolution and remarkable operative versatility. The echo signals obtained by ultrasonic inspection contain the information of flaws, based on which we can identify the flaw's position, magnitude and category. However, the conventional ultrasonic inspection techniques are based on manual or experiential pattern identification, which easily brings about costly, lengthy and erratic analysis. Therefore, a lot of modern signal processing methods and artificial intelligence tools have been applied to implement automatic ultrasonic signal classification [1]. In the automatic classification system, ultrasonic flaw signals acquired in a form of digitized data are first preprocessed, and then informative features are extracted and selected using various digital signal processing and pattern recognition techniques. Finally, a decision classifier can be trained for flaw identification, i.e., once the unseen ultrasonic signal features are inputted into classifier, the corresponding flaw type will be outputted.

In ultrasonic inspection process, the collected signals are presented in the format of an A-Scan, which is a plot of signal amplitude versus time [2]. Even though some flaw information can be obtained directly from the amplitude of the reflected signal, this time-amplitude presentation is not always the best representation of the signal. In most cases, the signal's frequency content is more important since some distinguished information is hidden in the frequency components. In order to find the frequency content of a signal, Fourier transform (FT) is often used for transforming the time based signal into frequency based one, i.e., the original signal is broken down into constituent sinusoids of different frequencies. The application of FT in ultrasonic inspection can be referred to [3-6].

For non-stationary ultrasonic signals, however, only the statistical parameters in time domain and frequency domain can be got by FT, while the local features about each flaw can not be extracted. Therefore, wavelet transforms (WT) based methods have been proposed to obtain more characteristic parameters for non-stationary signal. WT obtains the representation of a signal in terms of a finite length or fast decaying oscillating waveform, which is scaled and translated to match the input signal. In this way, it is possible to split local and global dynamics for a signal by

a multiresolution analysis (MRA) in a wavelet domain to solve the shortcoming of FT. WT based methods have gained wider application of feature extraction for ultrasonic flaw signals [7-10].

Not all features extracted from ultrasonic signals for a given classification problem need to be used due to their redundancy, therefore a further process is needed for redundancy reduction by retaining only small number of informative features. This stage of processing is called feature selection. Generally, two primary schemes are used for feature selection. One is to find a transformation from the original feature variables to a lower-dimensional feature space. The most widely used method for ultrasonic flaw signal classification is principal components analysis (PCA) [11-13]. Another is to select a subset from the original features by some criterions, which is described in [3, 14].

This study aims to present the optimal feature extraction and selection scheme for ultrasonic flaw signal classification. The main contributions of the study include: (1) Utilize empirical mode decomposition (EMD) to extract features for ultrasonic flaw signal. Unlike WT which needs pre-determined wavelet functions, EMD is more adaptive since it is based on signal information itself. (2) Apply rough set attribute reduction (RSAR) for further feature selection. Compared to PCA, RSAR can implement reduction, and in the meantime, preserve the original meaning of the features. The rest of this paper is organized as follows. Section 2 introduces the EMD method for feature extraction. Section 3 introduces the RSAR method for feature selection. Section 4 presents the artificial neural network (ANN) architecture for ultrasonic flaw signal classification. Section 5 and 6 describe the experimental procedure, result and analysis. Section 7 addresses the conclusions.

## II. EMPIRICAL MODE DECOMPOSITION FOR FEATURE EXTRACTION

### a. The basic principle of EMD

As mentioned in previous section, WT based methods have been widely used to extract features for ultrasonic flaw signal. The main advantage of such multiscale methods is to be suitable for signal analysis especially in highly noisy environments. For nonstationary ultrasonic signals, WT outperforms FT due to its ability to adapt the window size of the processed signal. Therefore, one can easily separate information and noise without needing a complex windowing step [15]. However, the main shortcoming of WT is that the time localization is poor for low frequency

signals and the frequency resolution is poor for high frequency signals. Moreover, the analysis result depends on the choice of the wavelet functions.

Empirical mode decomposition (EMD) is another powerful tool for adaptive multiscale analysis of nonstationary and nonlinear signals, which is similar with selective filter bank decompositions [16]. It is the construction of some intrinsic mode functions (IMFs) through a sifting process ended by an interpolation method and a stopping criterion. Compared to the original signal, IMF component is much simpler since it is a stationary narrow band signal charactering the information of original signal at different time scales. The small signal, easily inundated by big signal or background noise, can be clearly and effectively manifested in IMF component. IMF need satisfy two conditions: (1) At any instant time, the mean of upper envelope constituted by the local maxima and that of lower envelope constituted by the local minima are all zero. (2) In the whole data series, the number of extreme values (including the maxima and minima) and that of zero crossings are equal or differ at most by 1 [17].

For a given ultrasonic signal  $x(n)$ , we have the following equation by using EMD.

$$x(n) = \sum_{j=1}^p c_j(n) + r_p(n) \quad (1)$$

where  $c_j(n)$  ( $1 \leq j \leq p$ ) are the IMFs and  $r_p(n)$  is the residue. The process of EMD can be summarized as follows [18].

- (1) Identify all extrema of  $x(n)$ .
- (2) Interpolate between minima as well as maxima, ending up with some envelope  $e_{\min}(n)$  and  $e_{\max}(n)$  respectively.
- (3) Compute the average  $m(n) = (e_{\min}(n) + e_{\max}(n)) / 2$ .
- (4) Extract the residue  $r(n) = x(n) - m(n)$ . Repeat previous process until  $r(n)$  satisfies the two conditions mentioned above, and record it as IMF1.
- (5) Iterate above process and obtain a group of IMFs and a residue.

Note that among all IMFs, some are useful while some are irrelevant to flaw characteristic. We can select the representative IMFs for feature extraction by calculating the energy  $E(j)$  of each IMF, which is defined as

$$E(j) = \sum_{n=1}^N c_j^2(n) \quad (2)$$

where  $N$  is the total data length.

### b. Feature extraction based on EMD

Since it is very difficult to classify and recognize different flaw directly from the time domain features of the original ultrasonic signal, a series of IMFs of the original signal can be first obtained by using EMD. As mentioned 2.1, the IMFs contain useful information about flaw characteristic of original signal at different time scales, their corresponding traditional time and frequency domain based statistical parameters can be taken as the initial features for classification [19].

#### (1) Time domain based features

Since the IMFs are simple and stationary narrow band components of original signal, the statistical features can effectively reflect some characteristics of different flaws. In this study, a total of 10 time domain statistical features are extracted from the IMFs, which are described in table 1.

#### (2) Frequency domain based features

Note that the direct FT representation of original ultrasonic signal does not take into account any temporal information, therefore it can not effectively extract the flaw characteristic. While applying FT to IMFs of original signal, however, it can avoid such drawback by keeping both the temporal and frequency signal representation. In this study, 4 frequency domain statistical features are extracted, which are described in table 2. Note that  $x'(t)$  is the FT representation of  $x(t)$ .

## III. ROUGH SET ATTRIBUTE REDUCTION FOR FEATURE SELECTION

### a. Feature discretization by spectral clustering

Assume that the number of selected IMFs is  $m$ , thus a total of  $14m$  features can be extracted from each ultrasonic flaw signal. With the extracted features, different flaw signals can be identified and classified using pattern recognition tools such as ANN. Since high dimensional input feature vector may take more computational complexity for training ANN, feature selection scheme can be adopted to solve the problem. Feature selection aims to choose those features that are most predictive for a given outcome. Rough set attribute reduction (RSAR) is a representative method for this purpose. Unlike PCA method, RSAR can preserve the original meaning of the features after reduction. Furthermore, since the number of the total features in this study is not huge,

RSAR can implement feature selection in a concise way with high speed.

Table 1: Time domain features

Extracted feature	Value
Mean value $f_m$	$f_m = \frac{1}{n} \sum_{t=1}^n x(t)$
Variance $f_v$	$f_v = \frac{1}{n} \sum_{t=1}^n (x(t) - f_m)^2$
Root mean square $f_{rms}$	$f_{rms} = (\frac{1}{n} \sum_{t=1}^n x^2(t))^{1/2}$
Peak value $f_p$	$f_p = \max(x(t))$
Skewness $f_{ske}$	$f_{ske} = (\frac{1}{n} \sum_{t=1}^n (x(t) - f_m)^3) / f_v^{3/2}$
Kurtosis $f_{kur}$	$f_{kur} = (\frac{1}{n} \sum_{t=1}^n (x(t) - f_m)^4) / f_v^2$

Table 2: Frequency domain features

Extracted feature	Value
Mean value $f'_m$	$f'_m = \frac{1}{n} \sum_{t=1}^n x'(t)$
Variance $f'_v$	$f'_v = \frac{1}{n} \sum_{t=1}^n (x'(t) - f'_m)^2$
Peak value $f'_p$	$f'_p = \max(x'(t))$
Centre frequency $f_{fre}$	

While applying RSAR, the initial features can be taken as the attributes of an information system. Since the data type in initial feature set is mostly continuous, we need first divide values of each feature into some classes and substitute them with several discrete values, which help to implement the attribute reduction with rough set theory. A spectral clustering algorithm is used for discretization since it can achieve automatic selection of the cluster number. Given value set of a feature  $F = \{v_1, v_2, \dots, v_n\}$ , the spectral clustering algorithm can partition the  $n$  variables into  $j$  clusters by following steps.

(1) Construct the corresponding similarity matrix  $S = (s_{ij})_{n \times n}$  of  $F$ . If  $i=j$  then  $s_{ij}=0$ , otherwise

$$s_{ij} = \exp(-\|f_i - f_j\|^2 / (2\sigma)^2).$$

- (2) Compute the stochastic matrix  $P = D^{-1} \cdot S$  and obtain its eigensystems  $\{\lambda_i, u_i\}_{i=1}^n$ , where  $D$  is a diagonal matrix whose  $(i, i)$ -element is the sum of  $S$ 's  $i$ th row.
- (3) Select the optimal cluster number  $k$  if we have  $\arg \max_k (|\lambda_k - \lambda_{k+1}|)$ .
- (4) Construct the matrix  $Y$  by stacking in columns the first  $k$  eigenvectors of the generalized eigensystem  $Sx = \lambda Dx$ .
- (5) Treat each row of  $Y$  as a point in  $\mathfrak{R}^k$ , and partition all points into  $k$  clusters by using  $k$ -means algorithm.
- (6) Assign the original feature value  $v_i$  in  $F$  to cluster  $C_j$  if and only if row  $i$  of the matrix  $Y$  is assigned to cluster  $C_j$ , where  $1 \leq j \leq k$ .

After clustering, all feature values in  $F$  will be transformed to  $k$  discrete variables.

#### b. Feature selection based on RSAR

The whole initial feature vectors after discretization can be used to form an information system  $(U, A)$ , where  $U$  is a nonempty finite set of objects, called universe,  $A$  is a nonempty attribute set, including conditional attributes (each attribute stands for a feature) and decision attribute (each value stands for a flaw type), such that  $\forall a \in A, a : U \rightarrow V_a$ , where  $V_a$  is the set of values that attribute  $a$  may take. Let  $P \subseteq A$ ,  $\text{IND}(P)$  is an indiscernibility relation over  $U$ . For  $X \subseteq U$ ,  $X$  can be approximated using only the information contained within  $P$  by constructing the  $P$ -lower and  $P$ -upper approximations of  $X$  [20].

$$\underline{P}X = \{x \in U \mid [x]_P \subseteq X\} \quad (3)$$

$$\overline{P}X = \{x \in U \mid [x]_P \cap X \neq \emptyset\} \quad (4)$$

where  $[x]_P$  is the equivalence classes of  $\text{IND}(P)$ .

The tuple  $(\underline{P}X, \overline{P}X)$  is called a rough set. Let  $P$  and  $Q$  be the attribute sets inducing equivalence relations over  $U$ , then the positive, negative, and boundary regions can be respectively defined as follows:

$$\text{POS}_P(Q) = \bigcup_{x \in U/Q} \underline{P}X \quad (5)$$

$$\text{NEG}_P(Q) = U - \bigcup_{x \in U/Q} \overline{P}X \quad (6)$$

$$\text{BND}_P(Q) = \bigcup_{x \in U/Q} \overline{P}X - \bigcup_{x \in U/Q} \underline{P}X \quad (7)$$

For  $P, Q \subseteq A$ , it is said that  $Q$  depends on  $P$  in a degree  $k$  ( $0 \leq k \leq 1$ ), if

$$k = \gamma_P(Q) = \frac{|\text{POS}_P(Q)|}{|U|} \quad (8)$$

For given  $P$ ,  $Q$  and an attribute  $a \in P$ , the significance degree of  $a$  is defined as follow.

$$\sigma_P(Q, a) = \gamma_P(Q) - \gamma_{P-\{a\}}(Q) \quad (9)$$

A given dataset may have many attribute reductions, and the collection of all reductions is defined as follow.

$$R_{\text{all}} = \{X \mid X \subseteq C, \gamma_X(D) = \gamma_C(D), \text{ and } \forall a \in X, \gamma_{X-\{a\}}(D) \neq \gamma_X(D)\} \quad (10)$$

where  $C$  and  $D$  are the conditional attribute set and decision attribute set in  $A$ . In  $R_{\text{all}}$ , the reduction with the minimal cardinality, called the minimal reduction, can be defined as

$$R_{\text{min}} = \{X \mid X \in R_{\text{all}} \text{ and } \forall Y \in R_{\text{all}}, |X| \leq |Y|\} \quad (11)$$

In this study, the QuickReduct algorithm [21] will be used to obtain the minimal reduction for the initial features.

#### IV. BP NETWORKS FOR CLASSIFICATION

Feed-forward neural networks trained with the standard back-propagation algorithm (hereinafter referred as BP networks) are widely used for classifying ultrasonic flaw signals. BP networks are supervised networks requiring a desired response to be trained [22]. They learn to transform input data into a desired response, and have been shown to approximate the performance of optimal statistical classifiers in difficult problems. BP networks are flexible and nonlinear models consisting of a number of neurons organized into multiple layers. The complexity of the network can be changed by varying the number of layers and the number of neurons in each layer. Given enough hidden neurons and enough training data, it has been shown that BP networks can approximate virtually any function to any desired accuracy. BP networks are valuable tools in problems when one has little or no knowledge about the form of the relationship between input feature vectors and their corresponding outputs. Since the features in different domains of ultrasonic echo signal are not directly relative to the flaw types, BP networks will be trained in this study for classification.

Generally, BP network requires three or more layers of processing units: an input layer which accepts the input feature vectors used in the classification procedure, one or more hidden layers, and an output layer with one neuron per class. When a training data from an input pattern is



presented at the input layer, the network neurons perform calculations in the successive layers until an output value is computed at each of the output neurons. Finally, the output layer indicates which the appropriate class for the input data is. Every processing unit in one particular layer is usually connected to every neuron in the layer above and below. The connections carry weights which encapsulate the behavior of the network and are adjusted during training. In the forward pass stage, an input feature vector is presented to the network and its output is precisely the components of the input pattern. For successive layers the input to each neuron is the sum of the scalar products of the incoming vector components with their respective weights. In the back propagation stage, the output of neuron at output layer is then sent to all neurons in the following layer, which continues through all the layers of the network until the output layer is reached and the output vector is computed. Once training is finished, the BP network can predict the output for a new unseen input. Weights are then updated when actual output becomes available. The learning procedure includes training and update, and can be considered as a nonlinear estimation problem where the weights are to be estimated [23].

## V. EXPERIMENTAL PROCEDURE

We first acquired the ultrasonic A-scan signals from carbon fiber reinforced polymer (CFRP) specimens with different flaws. Then we decomposed the original ultrasonic signals by EMD, and extracted the time/frequency domain based statistical parameters of IMFs to construct the initial feature set. After using spectral clustering for feature value discretization, the QuickReduct algorithm was implemented on the discrete feature set to obtain the minimal reduction, and all attributes in the minimal reduction were taken as the input of ANN classifier (hereinafter referred as EMD\_RSAR scheme).

### (1) Signal acquisition

A PXU T227 digital flaw detector was used to send ultrasonic waves into the CFRP specimens with delamination, debonding and void through a transducer operating at the central frequency of 5 MHz. An echo was reflected back each time when the ultrasonic wave encountered a discontinuity in the propagation medium. The A-scan signals were digitised at a sampling frequency of 100 MHz and sample length of 512 using a Sonotek STR 8100 A/D board, and then stored in a personal computer (PC). The ultrasonic inspection system is shown in figure 1. The

collected signals are composed by 100 signals with no flaw, 100 signals with delamination, 100 signals with debonding and 100 signals with void. The four kinds of representative original signals are shown in figure 2.

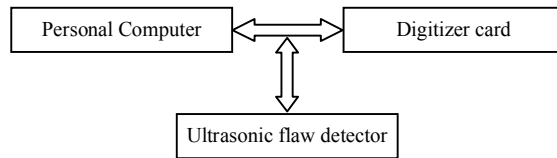


Figure 1. The ultrasonic inspection system

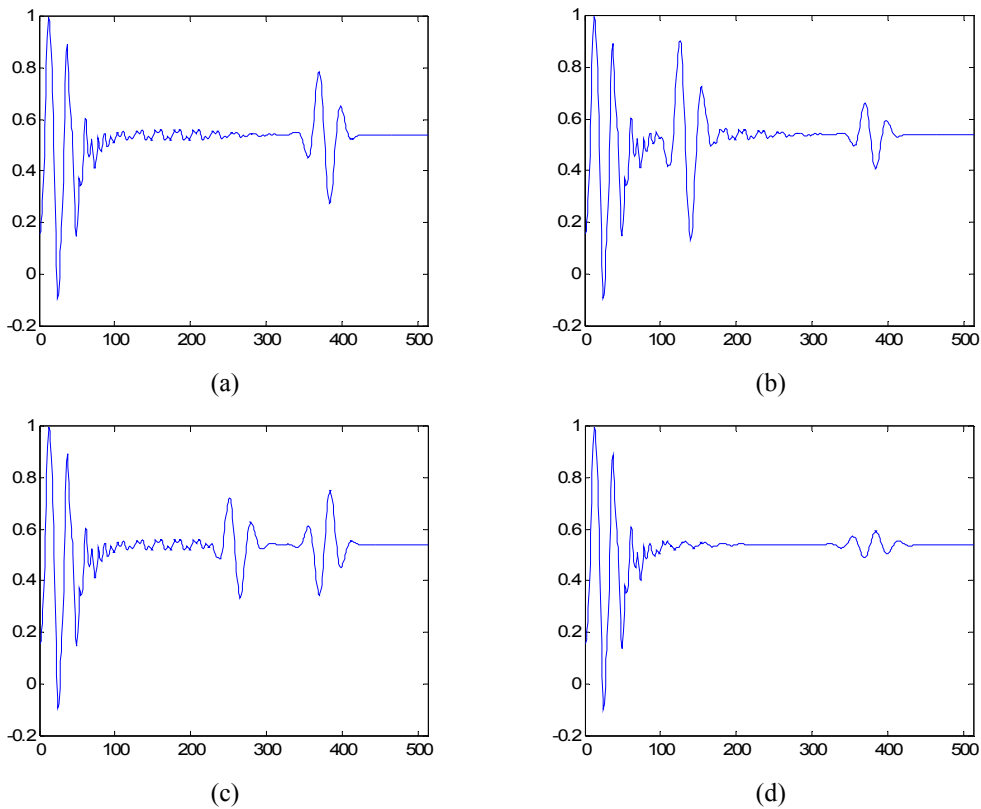


Figure 2. Four kinds of signals (a) No flaw (b) Delamination (c) Debonding (d) Void

## (2) Feature extraction

After signal acquisition, we obtained the IMFs of the different flaw signals by using EMD mentioned in section 2.1. The obtained IMFs are shown in figure 3.

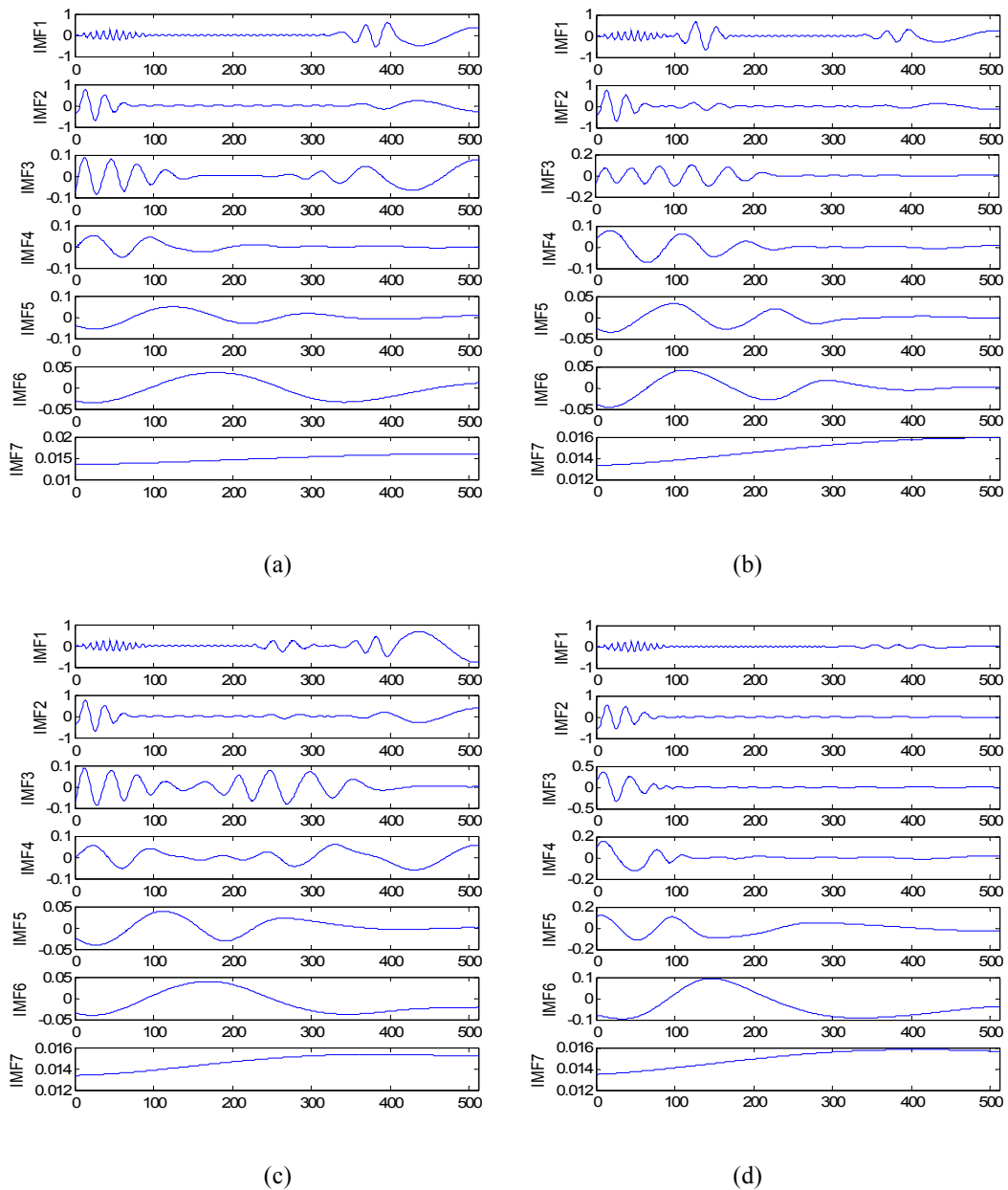


Figure 3. The IMFs of four kinds of signals (a) No flaw (b) Delamination (c) Debonding (d) Void

IMFs contain different aspects of original flaw signal, and some carry useful information while others may be relevant to noise. Eq. (2) was used to calculate the energy of each IMF. As is shown in figure 4, only the first five IMFs can be selected for further feature extraction because their energies are relatively high.

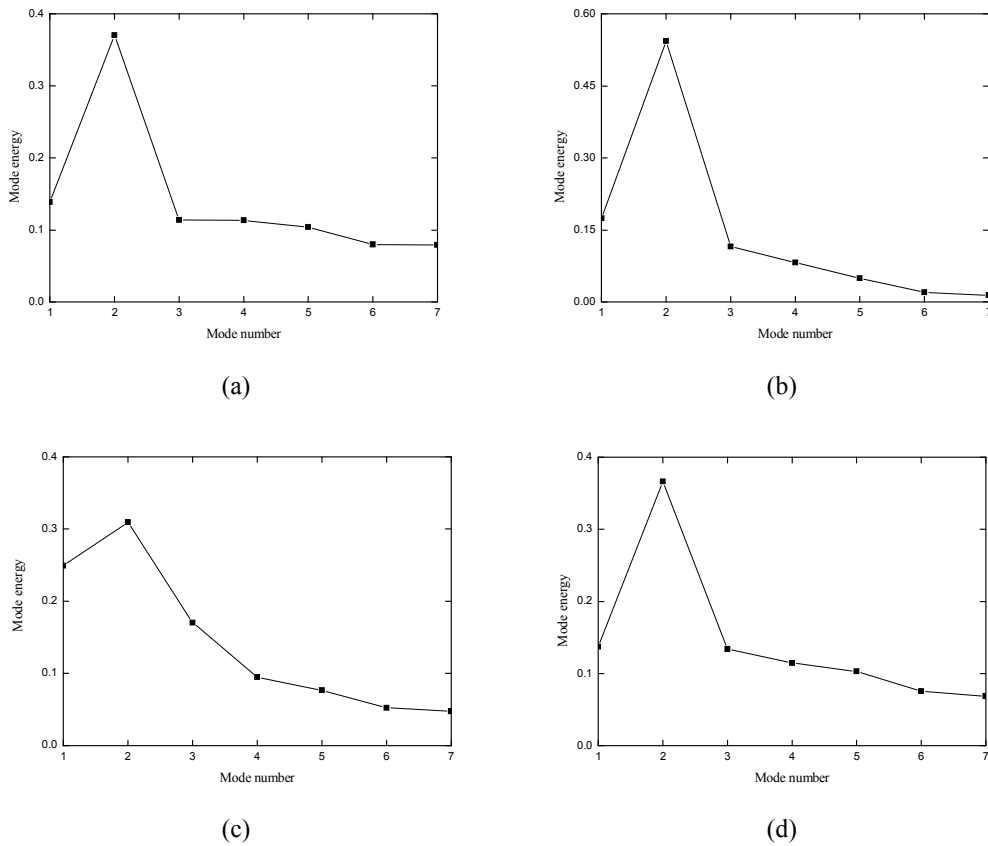


Figure 4. Energy of four kinds of signals (a) No flaw (b) Delamination (c) Debonding (d) Void

As mentioned in section 2.2, 14 time and frequency domain statistical features could be extracted from the each IMF. Thus, a total of 70 variables would be constructed the initial feature set for a given ultrasonic flaw signal. Figure 5 plots the FT representations of the IMF3 for four kind of flaw signal. Table 3 lists the 14 feature values of IMF3 of different flaw signals.

### (3) Feature selection

In this stage, the spectral clustering algorithm mentioned in section 3.1 was first used for feature value discretization. After that, we obtained 400 14-dimensional feature vectors, which could be formed an information system. Then the QuickReduct algorithm mentioned in section 3.2 was implemented on the information system to find the minimal reduction. The minimal reduction included 47 attributes (features), i.e.,  $f_m, f_v, f_p, f_{kur}, f_s, f_i, f_z, f'_m, f'_v, f_{fre}$  of IMF1,  $f_m, f_v, f_p, f_{ske}, f_s, f_z, f'_v, f'_p, f_{fre}$  of IMF2,  $f_m, f_v, f_p, f_{ske}, f_{kur}, f_s, f_i, f_z, f'_v, f_{fre}$  of IMF3,  $f_m, f_v, f_{rms}, f_{ske}, f_{kur}, f_s, f_z, f'_v, f_{fre}$  of IMF4, and  $f_m, f_v, f_p, f_{ske}, f_s, f_c, f_z,$

$f_v$ ,  $f_{fre}$  of IMF5, which would be taken as the final input vector of ANN. Note that it requires no additional parameters to operate other than  $\sigma$  ( $\sigma=50$  in this experiment) of the spectral clustering algorithm in the whole process of feature selection.

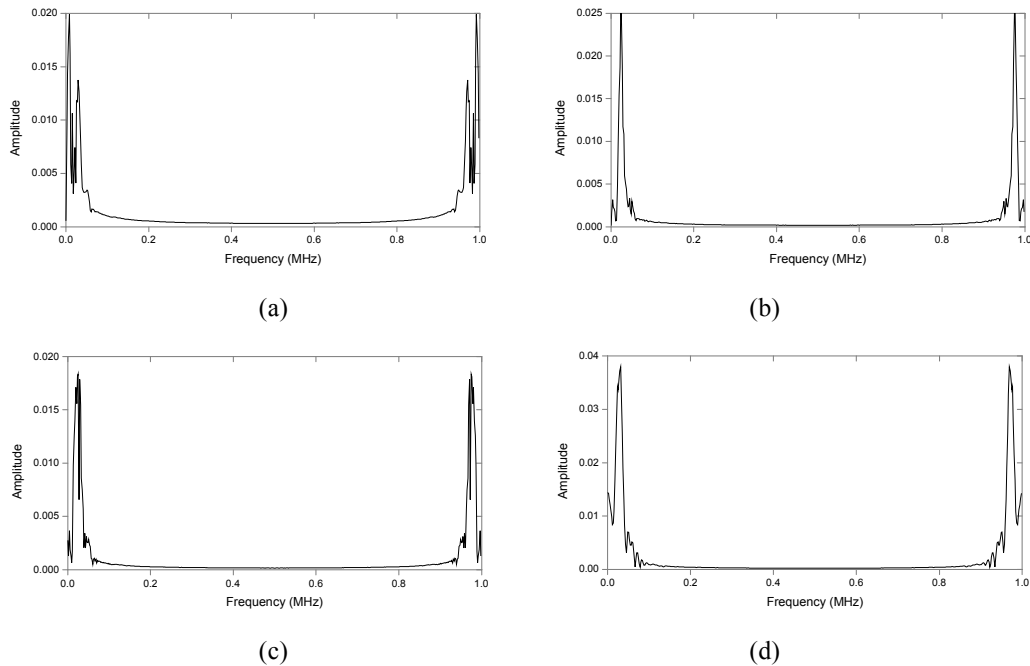


Figure 5. FT representations of IMF3 for four kinds of signals

(a) No flaw (b) Delamination (c) Debonding (d) Void

## VI. EXPERIMENTAL RESULTS AND ANALYSIS

The main purpose of the experiments is to investigate the effectiveness of the proposed EMD\_RSAR scheme for ultrasonic flaw signal classification. The other three feature extraction and selection schemes were also applied for comparison, which are listed as follows.

### (1) EMD\_PCA scheme

It first decomposed the original ultrasonic signal by EMD, and extracted the 14 time/frequency domain based statistical parameters of IMF1 to IMF5 (totally 70 variables) to construct the initial feature set. Then the principal components analysis (PCA) method was implemented on the whole feature sets, and the principal components (PCs) whose contributions to total variation of all 70 PCs are greater than 1% were selected to construct the final input vectors of ANN classifier.

The number of the selected PCs is 54 in this case.

Table 3: Features of IMF3 for different flaw signals (After normalization)

<b>Extracted feature</b>	<b>No flaw</b>	<b>Delamination</b>	<b>Debonding</b>	<b>Void</b>
$f_m$	0.03	0.01	0.16	0.81
$f_v$	0.13	0.14	0.13	0.60
$f_{rms}$	0.19	0.20	0.19	0.41
$f_p$	0.14	0.16	0.14	0.56
$f_{ske}$	0.07	0.03	0.04	0.92
$f_{kur}$	0.13	0.18	0.13	0.55
$f_s$	0.18	0.77	0.04	0.01
$f_c$	0.20	0.23	0.20	0.37
$f_i$	0.16	0.78	0.03	0.02
$f_z$	0.08	0.08	0.58	0.25
$f'_m$	0.23	0.18	0.18	0.41
$f'_v$	0.12	0.15	0.13	0.60
$f'_p$	0.20	0.25	0.18	0.37
$f'_{fre}$	0.09	0.53	0.29	0.09

### (2) DWT\_PCA scheme

It first decomposed the original ultrasonic signal by DWT on three level using Daubechies5 wavelet. The obtained 512 DWT coefficients were compressed into 256 samples by discarding the last 256 coefficients ( $cD_1$ ), which do not contain much information but mainly noise, and constructed the initial feature set. Then the PCA method was implemented on the whole feature sets, and the PCs whose contributions to total variation of all 256 PCs are greater than 1% were selected to construct the final input vectors of ANN classifier. The number of selected PCs is 97 in this case.

### (3) Clustered\_DWT scheme

It first decomposed the original ultrasonic signal by DWT on three level using Daubechies5 wavelet. The obtained 512 DWT coefficients were compressed into 256 samples by discarding the last 256 coefficients ( $cD_1$ ), which do not contain much information but mainly noise, and constructed the initial feature set. Then the feature sets were further mapped to cluster energy features and stored as clustered\_DWT feature sets as mentioned in [1]. The number of clusters is set to 64 in this case.

BP networks with one hidden layer were used for classifying the ultrasonic signals into no flaw, delamination, void or debonding. To compare the four kinds of feature extraction and selection schemes, i.e., EMD\_RSAR, EMD\_PCA, DWT\_PCA and Clustered\_DWT, four ANN architectures respectively having 47, 54, 97 and 64 input nodes were designed. Kolmogorov's theorem was used for determining the number of neurons at hidden layer. The learning rate was set to 0.2 and the topological order was applied as the update mode of the networks. The 5-fold cross-validation was carried out for assessing classification performance of all ANNs. The 400 ultrasonic signals were shuffled and randomly divided up into 5 subsets. Note that the number of each kind of flaw signal is 20 in all subsets. In turn, 4 of these subsets were used to train the network, and the remaining subset was used to validate the network. The process did not terminate until every subset was taken as training set and test set. Moreover, we got a average of the network training ability by assigning 10 different initial weights to the network. The classification performance with different feature extraction and selection schemes could be compared using the result of each cross-validation test. The mean square error (MSE) limit was set to 0.001 for stopping the training process, and the epoch limit was set to 200,000 for those occasional cases where training failed to converge. The values of main parameters for training ANNs are resumed in table 4.

Table 4: The parameters of ANNs

<b>Parameter</b>	<b>Value</b>
No. of neurons at input layer	54/47/97/64
No. of neurons at output layer	4
No. of neurons at hidden layer	109/95/195/129
Activation function at hidden layer	tansig
Activation function at output layer	tansig
Training algorithm	trainlm
Performance goal	0.001

The average training elapsed times of ANN classifiers by using EMD\_PCA, EMD\_RSAR, DWT\_PCA and clustered\_DWT scheme were 58.8s, 48.7s, 105.6s and 60.7s respectively, which showed the efficiency of EMD\_RSAR scheme since it only selected 47 features for training ANN.

The classification accuracy in percentage of ANN classifiers by using different feature extraction

and selection schemes are summarized in table 5 to 8. Note that only one cross-validation training classification result was shown here due to the limit of space. Let us first analysis the performance of EMD\_PCA and EMD\_RSAR for classification. As are shown in table 5 and 6, the average recognition rate of EMD\_RSAR (90.94%) is higher than that of EMD\_PCA (89.38%). It validates that rough set attribute reduction can retain the most informative attributes (features) while reducing the amount of redundant attributes involved for classification. Since the features extracted from IMFs are highly relevant, it degrades the performance of PCA for feature selection. Moreover, to achieve higher computational efficiency, EMD\_PCA selects only 54 PCs as the input vector of ANN, which reduces the recognition rate of classification.

Then we focus on the performance analysis of EMD\_RSAR, DWT\_PCA and clustered\_DWT schemes for ultrasonic flaw signal classification. As are shown in table 6, 7 and 8, DWT\_PCA scheme achieves the worst classification performance, whose average recognition rate is only 87.5%. Obviously, DWT coefficients can not effectively describe the flaw characteristic because of the time-variance problem of ultrasonic signals. Clustered\_DWT scheme uses the unsupervised learning method to cluster the DWT coefficients on the basis of the average energies at different scales and within the same scale, which can achieve higher average recognition rate (89.06%). However, the performance of clustered\_DWT scheme still depends on the pre-determined wavelet basis function, and wavelet basis can not be changed once it is established, which can not guarantee the optimal decomposition. EMD\_RSAR achieves the highest classification performance (90.94%) since it can decompose time-variance signal into a series of IMFs, which represent real physical information and local characteristic of flaw signal. The limited number of IMFs guaranties that EMD is selfadaptive and fit for nonlinear and nonstationary signal processing. Furthermore, the optimal feature selection by rough set attribute reduction also improves the classification efficiency. Additionally, the number of false positive or negative using EMD\_RSAR for classification is the least. Especially for the false negative (see table 6), only one void signal was classified as no flaw by using EMD\_RSAR scheme, which is greatly lower than the rate by using DWT\_PCA or clustered\_DWT. It is extraordinary important in some applications where undetected flaw are more risky compared to false alarms.

Table 5: Classification result using EMD\_PCA based features



Class	Number of training samples	Number of samples assigned into this class				Recognition rate (%)
		No flaw	Delamination	Debonding	Void	
No flaw	80	73	1	2	4	91.25
Delamination	80	0	70	7	3	87.5
Debonding	80	1	3	69	7	86.25
Void	80	1	1	4	74	92.5

Table 6: Classification result using EMD\_RSAR based features

Class	Number of training samples	Number of samples assigned into this class				Recognition rate (%)
		No flaw	Delamination	Debonding	Void	
No flaw	80	75	0	2	3	93.75
Delamination	80	0	72	6	2	90
Debonding	80	0	3	70	7	88.63
Void	80	1	1	4	74	92.5

Table 7: Classification result using DWT\_PCA based features

Class	Number of training samples	Number of samples assigned into this class				Recognition rate (%)
		No flaw	Delamination	Debonding	Void	
No flaw	80	73	1	2	4	91.25
Delamination	80	1	69	6	4	86.25
Debonding	80	2	4	66	8	82.5
Void	80	1	2	5	72	90

Table 8: Classification result using clustered\_DWT based features

Class	Number of training samples	Number of samples assigned into this class				Recognition rate (%)
		No flaw	Delamination	Debonding	Void	
No flaw	80	74	0	2	4	92.5
Delamination	80	0	70	7	3	87.5
Debonding	80	1	4	68	7	85
Void	80	1	2	4	73	91.25

## VII. CONCLUSIONS

This paper presented a EMD\_RSAR scheme to extract and select features for ultrasonic flaw signal classification. It first decomposed the original ultrasonic flaw signal into a series of IMFs by using EMD, and extracted their corresponding traditional time and frequency domain based

statistical parameters as the initial features. Then it utilized spectral clustering algorithm to divide values of each feature into some classes and substitute them with several discrete values. After that, the rough set attribute reduction method was applied for implementing the feature selection. Finally, BP networks were trained by input features obtained by different schemes, including EMD\_PCA, EMD\_RSAR, DWT\_PCA and clustered\_DWT. Experimental results shows that EMD\_RSAR could achieve high classification accuracy and training efficiency with lower false positive as well as false negative, which is a promising feature extraction and selection scheme for ultrasonic flaw signal identification.

#### ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of JiangXi Province (No. 20122BAB201039), the Foundation of Key Laboratory of Nondestructive Testing (Nanchang Hangkong University), Ministry of Education (No. ZD201229003), the Scientific Research Foundation of Education Department of JiangXi Province (No. GJJ13515), the Scientific Research Foundation of Chongqing Municipal Education Commission (No. KJ111218).

#### REFERENCES

- [1] K. Lee and V. Estivill Castro, "Feature extraction and gating techniques for ultrasonic shaft signal classification", *Applied Soft Computing*, Vol. 7, No. 1, pp. 156-165, 2007.
- [2] Kyungmi Lee, "Feature Extraction Schemes for Ultrasonic Non-destructive Testing Inspections", *Advances in Information Sciences and Service Sciences*, Vol. 3, No. 3, pp. 125-135, 2011.
- [3] A.A. Anastassopoulos, V.N. Nikolaidis and T.P. Philippidis, "A Comparative Study of Pattern Recognition Algorithms for Classification of Ultrasonic Signals", *Neural Computing & Applications*, Vol. 8, No. 1, pp. 53-66, 1999.
- [4] A. Shelke, T. Kundu, U. Amjad, K. Hahn and W. Grill, "Mode-selective excitation and detection of ultrasonic guided waves for delamination detection in laminated aluminum plates", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 58, No. 3, pp. 567-577, 2011.

- [5] K. Lee, "A Hybrid Classification Approach to Ultrasonic Shaft Signals", Lecture Notes in Computer Science, Vol. 3339, pp. 11-17, 2005.
- [6] K. Lee, "Feature extraction schemes for ultrasonic signal processing", Proc. IEEE 5th International Conference on Computer Sciences and Convergence Information Technology, pp. 366-372, 2010.
- [7] M. Cacciola, S. Calcagno and F.C. Morabito, "Computational intelligence aspects for defect classification in aeronautic composites by using ultrasonic pulses", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 55, No. 4, pp. 870-878, 2008.
- [8] Zhao Chihang, Zhong Xin, Dang Qian and Zhao Liye, "De-noising Signal of the Quartz Flexural Accelerometer by Multiwavelet Shrinkage", International Journal on Smart Sensing and Intelligent Systems, Vol. 6, No. 1, pp. 191-208, 2013.
- [9] S. Sambath, P. Nagaraj and N. Selvakumar, "Automatic Defect Classification in Ultrasonic NDT Using Artificial Intelligence", Journal of nondestructive evaluation, Vol. 30, No. 1, pp. 20-28, 2011.
- [10] H. Hashim, S. Ramli, N. Wahi, M. S. Sulaiman and N. Hassan, "Recognition of Psoriasis Features via Daubechies D8 Wavelet Technique", International Journal on Smart Sensing and Intelligent Systems, Vol. 6, No. 2, pp. 711-732, 2013.
- [11] G.P.P. Gunarathne and Y. Qureshi, "Development of a synthetic A-scan technique for ultrasonic testing of pipelines", IEEE Transactions on Instrumentation and Measurement, Vol. 54, No. 1, pp. 192-199, 2005.
- [12] U.A. Qidwai, "Autonomous corrosion detection in gas pipelines: a hybrid-fuzzy classifier approach using ultrasonic nondestructive evaluation protocols", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 56, No. 12, pp. 2650-2665, 2009.
- [13] N. Bochud, A.M. Gomez and G. Rus, "Robust parametrization for non-destructive evaluation of composites using ultrasonic signals", Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1789-1792, 2011.
- [14] X.J. Zhou and T.S. Dillon, "A statistical-heuristic feature selection criterion for decision tree induction", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 8, pp. 834-841, 1991.

- [15] Sylvie Legendre, Daniel Massicotte, Jacques Goyette and Tapan K. Bose, "Wavelet Transform Based Method of Analysis for Lamb-Wave Ultrasonic NDE Signals", IEEE Transactions on Instrumentation and Measurement, Vol. 49, No. 3, pp. 524-530, 2000.
- [16] C. Damerval and S. Meignen, "A Fast Algorithm for Bidimensional EMD", IEEE Signal Processing Letters, Vol. 12, No. 10, pp. 701-704, 2005.
- [17] Yanhua Zhang, Lu Yang and Jianping Fan, "Study on feature extraction and classification of ultrasonic flaw signals", WSEAS Transactions on Mathematics, Vol. 9, No. 7, pp. 529-538, 2010.
- [18] T. Chen, P. Que and Q. Zhang, "Ultrasonic signal identification by empirical mode decomposition and Hilbert transform", Review of scientific instruments, Vol. 76, No. 8, pp. 1-6, 2005.
- [19] Kiran George and Chien-In Henry Chen, "Biologically-Inspired Signal Processor Using Lateral Inhibition and Integrative Function Mechanisms for High Instantaneous Dynamic Range", International Journal On Smart Sensing and Intelligent Systems, vol. 4, no. 4, pp. 547-567, 2011.
- [20] Richard Jensen and Qiang Shen, "New Approaches to Fuzzy-Rough Feature Selection", IEEE Transactions on Fuzzy Systems, Vol. 17, No. 4, pp. 824-838, 2009.
- [21] Ye Hongxian, Li Wenchang and Hu Xiaoping, "A Blind Source Separation Method for Convolved Mixtures by Non-Stationary Vibration Signals", International Journal On Smart Sensing and Intelligent Systems, Vol. 6, No. 3, pp. 973-992, 2013.
- [22] S. Iyer, S.K. Sinha and B.R. Tittmann, "Ultrasonic signal processing methods for detection of defects in concrete pipes", Automation in Construction, Vol. 22, pp. 135-148, 2011.
- [23] Zhao Chihang, Zhong Xin, Dang Qian, Zhao Liye, "De-noising Signal of the Quartz Flexural Accelerometer by Multiwavelet Shrinkage", International Journal On Smart Sensing and Intelligent Systems, Vol. 6, No. 1, pp. 191-208, 2013.