Application of EEMD Sample Entropy and Grey Relation Degree in Gearbox Fault Identification

Wenbin Zhang*, Libin Yu, Yanping Su, Jie Min and Yasong Pu

College of Engineering, Honghe University, Mengzi 661100, China

Abstract: In this paper, a new gearbox fault identification method was proposed based on mathematical morphological filter, ensemble empirical mode decomposition (EEMD), sample entropy and grey relation degree. Firstly, the sampled data was de-noised by mathematical morphological filter. Secondly, the de-noised signal was decomposed into a finite number of stationary intrinsic mode functions (IMFs) by EEMD method. Thirdly, some IMFs containing the most dominant fault information were calculated by the sample entropy for four gearbox conditions. Finally, since the grey relation degree has good classified capacity for small sample pattern identification, the grey relation degree between the symptom set and standard fault set was calculated as the identification evidence for fault diagnosis. The practical results show that this method is quite effective in gearbox fault diagnosis. It's suitable for on-line monitoring and fault diagnosis of gearbox.

Keywords: Ensemble empirical mode decomposition, Feature extraction, Gearbox, Grey relation degree, Identification, mathematical morphological filter, Sample entropy.

1. INTRODUCTION

Gear is a key component usually used in mechanical transmission, for its prominent carrying capacity and reliability. Therefore, the fault identification of gear has been the subject of extensive research. Enveloping analysis and wavelet package decomposition are commonly used in fault diagnosis as feature extraction methods for gear signal [1]. But the enveloping analysis needs to confirm the center frequency and frequency band of band-pass filter, and it will impact the analytical results [2]. While the wavelet decomposition has finite length of basic function, energy will leak in the signal processing. Because the wavelet decomposition is based on the linear decomposition, the good effectiveness will not be obtained in gear fault data processing due to its non-linear and non-stationary behaviors.

Mathematical morphology is a subject concerned with the shape of an object based on set theory and integral geometry [3]. In recent years, more and more studies have been done on the morphological filter. It is a non-linear filter with the advantage of better performance on rejection of white noise and pulse noise [3]. The operations of the filter are mainly plus, minus and logic. So the implementation of the filter by software or hardware is very easy. Its filtering idea is based on the geometrical structure of the filtered signals and realized through moving predefined structure element to match and adjust the singular parts of the signals [4]. It has been used in signal de-noising and purification of rotor axis [5, 6].

Gear fault signal is the typical non-stationary and non-linear signal. How to extract feature parameter of different fault patterns is the key for gear fault diagnosis. Sample entropy is a good tool to evaluate complexity of non-linear time series, compared with other existing non-linear dynamic methods. It has many good characteristics, such as good resistance of noise interference and closer agreement with theory for data sets with known probabilistic content. Moreover, sample entropy displays the property of relative consistency in some situations where approximate entropy does not [7]. These performances are suitable for fault extraction in practice.

In order to extract fault feature of gearbox, in this paper, a novel approach is proposed based on mathematical morphological filter, EEMD and grey relation degree. The proposed method could extract gear fault feature by EEMD and sample entropy. Then we identify a different gearbox fault mode by calculating the grey relation degree between the fault sample and standard fault pattern.

2. BASIC CONCEPTS OF MATHEMATICAL MORPHOLOGICAL FILTER

A mathematical morphological filter is constructed by different morphological transformations. First, several important morphological transformations are introduced.

Dilation and erosion are two basic morphological transformations. While dilation is the transformation used to expand the targeted object and shrink the hole, erosion is the transformation used to shrink the targeted object and expand the hole. Let $f(x)$ and $g(x)$ denote one-dimensional input signal and structure element, where $F=\{0, 1, \ldots, N-1\}$ and $G=\{0, 1, \ldots, M-1\}$ denote sets in which signal f and g are defined, here $N \geq M$. Dilation and erosion of f and g are thus defined as follows:

$$(f \oplus g)(n) = \max_{m=0,1,\cdots,M-1} \{f(n-m)+g(m)\}$$

*Address correspondence to this author at the College of Engineering, Honghe University, Jinhua Road, Mengzi, 661100, China;
Tel: 15025218982; E-mail: 190322507@qq.com

1874-155X/14 2014 Bentham Open

$(n=0, 1, \ldots, N-M)$ (1)

$$(f\Theta g)(n) = \min_{m=0,1,\cdots,M-1} \{f(n+m) - g(m)\}$$

$(n=0, 1, \ldots, N+M-2)$ (2)

Usually, dilation and erosion are not mutually inverse. They can be combined through cascade connection to form new transformations. If dilation is next to erosion, such cascade transform is an opening transformation. The contrary is a closing transformation. The transformations can be computed using the following formulae respectively

$$(f \circ g)(n) = [f \Theta g \oplus g](n) \quad (3)$$

$$(f \bullet g)(n) = [f \oplus g \Theta g](n) \quad (4)$$

The opening and closing results of the signal f by the elliptical structure element g are shown in Fig. (**1**) [8].

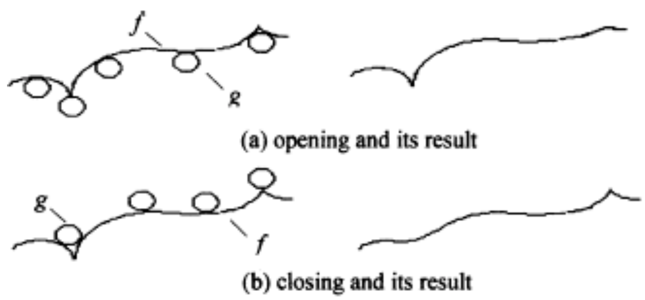


Fig. (1). Opening and closing results of the signal f by the elliptical structure element g.

From Fig. (**1a**), we can know that when g moves under f closely, the parts of f that do not contact with g will fall into the upper edge of g. So the opening transformation can be used to remove the peaks in the signal. From Fig. (**1b**), when g moves over f closely, the parts of f that do not contact with g will roll into the lower edge of g. So the closing transformation can be used to fill the valleys in the signal. Both transformations can be combined to form a morphological filter because they have the capacity of low-pass filtering.

In order to reject both positive and negative noise together, the open-closing filter and close-opening filter can be realized with the same structure element through cascade connection of the opening and closing transformation in different order. Two filers are defined as follows:

$$OC(f(n)) = (f \circ g \bullet g)(n) \quad (5)$$

$$CO(f(n)) = (f \bullet g \circ g)(n) \quad (6)$$

Due to the expansibility of the opening transformation and the inverse expansibility of the closing transformation, the problem of the statistics deviation exists in the open-closing filter and close-opening filter. The output of the open-closing filter is small, while that of the close-opening filter is large. Under most circumstances, the best processing performance can't be achieved by using a single filter. In order to lower the output deviation, two filters can be cascaded to form a new combined filter, whose output is defined as the following:

$$y(n) = [OC(f(n)) + CO(f(n))]/2 \quad (7)$$

Actually, the structure element acts as a filtering window, in which the data are smoothed to have a similar morphological structure as the structure element. The effectiveness and accuracy of the morphological filter depend on not only the combination mode of different transformations, but also the shape and width of the structure element. Usually the shape of the structure element should be similar to the signal. The commonly used structure element has simple geometrical shape, such as line, round, triangle and other polygon etc. In general, the more complex the structure element, the better the effectiveness will be obtained to reject the noises, but it will cost much time.

For structure element with the assured shape, it is necessary to select proper height and length, of which the length is especially more important to the effectiveness of signal processing. In vibration signal processing, the selection of the height is based on the experience. For the triangular structure element, selecting 1 to 5 percent of original signal's height is appropriate [9]. The length is mainly determined by the period and sampling frequency of signal's main wave. Meanwhile, only if the width of the structure element is longer than that of the widest pulse in the series, can all the pulse interferences be removed.

3. BASIC CONCEPT OF EEMD

The concept of the EEMD is the following: the added white noise consists of components with different scale and would be uniform to inhabit the whole time-frequency space. When a signal is added with the uniformly distributed white noise background, the different scale components of the signal are automatically projected onto proper scales of reference established by the white noise in the background. Because each of noise-added decompositions contains the signal and the added white noise, each individual trial is certain to get very noisy results. As the noise in each trial is different from separate trials, the noise can be almost completely removed by the ensemble mean of entire trials. The ensemble mean is treated as the true answer because only the signal is persevered finally as more and more trials are added in the ensemble. The crucial principle advanced here is based on the following observations [10, 11]:

(1) A collection of white noise cancels each other out in a time-frequency ensemble mean; therefore, only the signal can continue to exist and remain in the final noise-added ensemble mean.

(2) White noise of finite amplitude necessarily compels the ensemble to discover all possible solutions. The white noise makes the different scale signals reside in the corresponding IMFs, controlled by dyadic filter banks, and renders the results of ensemble mean more meaningful.

(3) The decomposition result with truly physical meaning of the empirical mode decomposition is not the one without noise; it is assigned to be the ensemble mean of a large number of trials comprising the noise-added signal.

Based on the aforementioned observations, the EEMD algorithm can be stated as follows [10, 12]:

(1) Initialize the ensemble number M and the amplitude of the added white noise, let M=1.

(2) Execute the mth trial for the signal added white noise.

(a) Add the white noise series with the given amplitude to the investigated signal, i.e.

$x_m(t)=x(t)+n_m(t)$ (8)

where $n_m(t)$ represents the mth added white noise, and $x_m(t)$ indicates the noise-added signal of the mth trial.

(b) Decompose the noise-added signal $x_m(t)$ into l IMFs $c_{i,m}(i=1,2,\ldots l,\ m=1,2,\ldots,M)$ using the empirical mode decomposition method. Where $c_{i,m}$ indicates the ith IMF of the mth trial; l is the number of IMFs and M means the number of the ensemble.

(c) If $m<M$, then let $m=m+1$ and repeat the step (a) and (b) again and again until $m=M$, but with different white noise each time.

(3) Compute the ensemble mean $\bar{c}_i$ of the M trials for each IMF, and we obtain

$$\bar{c}=\frac{1}{M}\sum c \tag{9}$$

(4) Report the mean $\bar{c}_i$ ($i=1,2,\ldots,l$) of each of l IMFs as the final ith IMF.

4. DEFINITION OF SAMPLE ENTROPY

Let $[x(n)]=x(1), x(2), \ldots, x(N)$ denote N-dimensional elements of time series representing rotor vibration signal. Then, the estimation algorithm of sample entropy consists of the following steps [13]:

(i) Creating m vectors is defined as:

$$X_m(i)=[x(i), x(i+1), \ldots, x(i+m-1)].\ (i=1, 2, \ldots, N-m+1) \tag{10}$$

(ii) Calculation of distance between two vectors in the following way:

$$d[X_m(i), X_m(j)]=\max_{k=0,\cdots,m-1}|x(i+k)-x(j+k)| \tag{11}$$

(iii) Calculation of number of similar segments in two vectors:

$n_m=\#\ d[X_m(i), X_m(j)]\le r$, while $i\neq j$

$n_{m+1}=\#\ d[X_{m+1}(i), X_{m+1}(j)]\le r$, while $i\neq j$

where, r is a tolerance parameter.

(iv) Calculation of similarity measures of these segments:

$$B_i^m(r)=\frac{1}{N-m+1}n_m \quad A_i^m(r)=\frac{1}{N-m+1}n_{m+1}$$

$i=1,\ldots,N-m$.

(v) Calculation of mean measures of the similar signal segments:

$$B_m=\frac{\sum_{i=1}^{N-m}B_i^m(r)}{N-m} \quad A_m=\frac{\sum_{i=1}^{N-m}A_i^m(r)}{N-m}$$

(vi) Calculation of sample entropy estimation:

$$\text{SampEn}(m,r)=-\ln\frac{A^m(r)}{B^m(r)} \tag{12}$$

5. GREY RELATION DEGREE

According to the grey theory, the relation degree evolves from the relation coefficient. The relation coefficient of the two series X_i and X_j, is represented by $\zeta_{ij}(k)$, where k represents the sampling points [14, 15].

$$\Delta_{ij}(k)=|X_j(k)-X_i(k)|\ k\in\{1,2,\ldots,N\} \tag{13}$$

$$\Delta\text{min}=\min_j\min_k\Delta_{ij}(k)\quad \Delta\text{max}=\max_j\max_k\Delta_{ij}(k)$$

$\zeta_{ij}(k)$ is defined as:

$$\xi_{ij}=\frac{\Delta_{\min}+\Delta_{\max}\cdot\rho}{\Delta_{ij}(k)+\Delta_{\max}\cdot\rho}\quad k\in\{1,2,\ldots,N\} \tag{14}$$

where ρ is a constant with the range from 0 to 1. The value of ρ determines the classification capacity and is usually recommended to be 0.5. The relation degree of the two series X_i and X_j is as follows:

$$\xi_{ij}=\frac{1}{N-1}\cdot\frac{1}{2}[\sum_{k=1}^{N}\xi_{ij}(k)+\sum_{k=2}^{N-1}\xi_{ij}(k)] \tag{15}$$

The relation degree represented by ζ_{ij} shows the comparability of the X_i and X_j series. It is often applied to grey cluster in practice [16]. Obviously, the bigger the ζ_{ij} is, the greater the inference of X_i to X_j would be.

6. ALGORITHM OF FAULT IDENTIFICATION OF GEARBOX

The detailed algorithm of fault diagnosis can be seen as below:

Step 1: The sample data are obtained from the experimental testing of gearbox under four conditions, which are normal, slight-worn, medium-worn and broken-teeth.

Step 2: Using mathematical morphological filter to de-noise the white noise and other interferences in the original signal.

Step 3: Using EEMD to process the de-noised signals. Select some IMFs which contain the most dominant fault information as research objects.

Step 4: Calculating sample entropy of these IMFs by equation (12).

Step 5: Building the feature vector by equation (16):

$$[T]=[SE_1, SE_2, \ldots, SE_i] \tag{16}$$

Where i refers to the number of selected IMFs.

Step 6: The grey relation degree between the symptom set and standard fault set is calculated as the identification evidence.

7. PRACTICAL APPLICATION

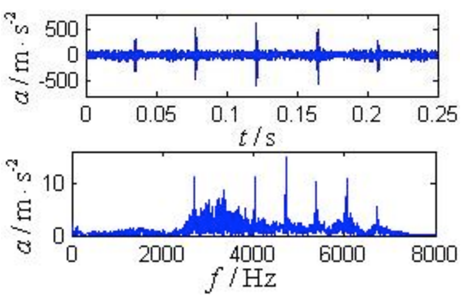
To verify good effectiveness in gearbox fault identification, all vibration signals were collected from the experimental testing of gearbox using the accelerometer which was mounted on the outer surface of the bearing case

Table 1. Sample entropy of different gearbox fault pattern.

Gearbox Condition	SE_1	SE_2	SE_3	SE_4	SE_5
Normal	0.8247	0.6391	0.5761	0.3630	0.1611
Broken-teeth	0.6688	0.5566	0.4976	0.2666	0.1498
Slight-worn	0.9748	0.6352	0.5815	0.2778	0.1748
Medium-worn	0.8821	0.6319	0.5715	0.3231	0.1772

of input shaft of the gearbox. The speed of the motor was 1420 RPM and the sample frequency was 16384 Hz. Testing were carried out under four kinds of conditions which were normal, slight-worn, medium-worn and broken-teeth. We got five sets of sampled data under each condition. First, we used mathematical morphological filter to process the original signal. Fig. (**2a**) shows the waveform of the original broken-teeth signal in time and frequency domain. Fig. (**2b**) shows the processed signal.

(a) Original broken-teeth signal and its spectrum



(b) Processed signal and its spectrum

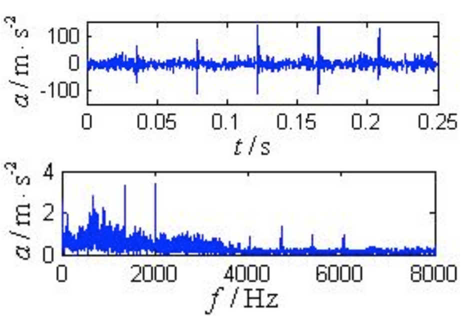


Fig. (2). Waveform of original signal and de-noised results comparison.

Comparing the above two figures, we can see that the high frequency noises are eliminated and the fault feature is obtained. It is very useful for the next procedure.

Next, we use EEMD to decompose the same signal. Fig. (**3**) gives the processed results.

From the above figure, we can see that de-noised vibration signals are decomposed into a finite number of stationary intrinsic mode functions (IMFs); and IMF 1 to IMF 5 contain obvious shocking components. So we calculate the sample entropy of these IMFs. Table **1** gives the mean calculated values of ten data sets in four fault conditions. From Table **1**, we can see that different fault mode has different sample entropy.

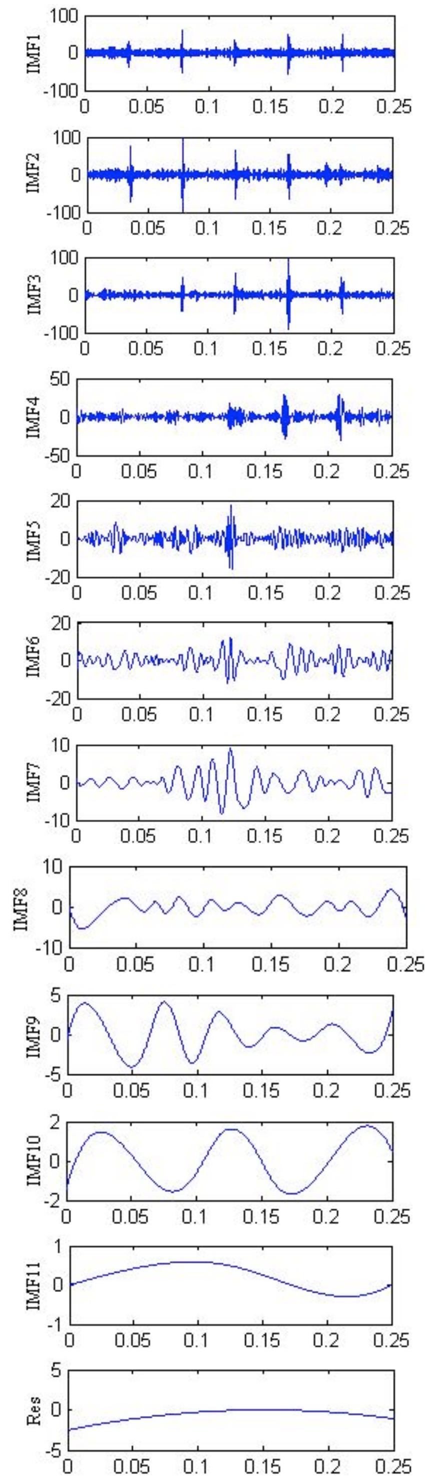


Fig. (3). EEMD decomposition results of broken-teeth signal.

Table 2. Sample Entropy extracted by EEMD in different conditions.

Gearbox Condition	Sample	SE_1	SE_2	SE_3	SE_4	SE_5
Normal	1	0.8269	0.6320	0.5827	0.3609	0.1458
	2	0.8146	0.6398	0.5813	0.3546	0.1614
	3	0.8138	0.6355	0.5633	0.3511	0.1501
	4	0.8379	0.6445	0.5881	0.3889	0.1585
	5	0.8413	0.6412	0.5820	0.3993	0.1726
Slight-worn	1	0.9903	0.6321	0.5776	0.2581	0.1771
	2	0.9852	0.6363	0.5877	0.2788	0.1700
	3	1.0139	0.6369	0.5827	0.2675	0.1759
	4	0.9545	0.6459	0.5816	0.2913	0.1821
	5	0.9725	0.6211	0.5882	0.2673	0.1705
Medium-worn	1	0.8719	0.6391	0.5688	0.2915	0.1832
	2	0.8577	0.6318	0.5698	0.3308	0.1865
	3	0.8942	0.6343	0.5732	0.3140	0.1752
	4	0.8890	0.6282	0.5645	0.3372	0.1741
	5	0.8966	0.6276	0.5693	0.3187	0.1781
Broken-teeth	1	0.6630	0.6145	0.4976	0.2259	0.1753
	2	0.6717	0.5580	0.4846	0.2836	0.1343
	3	0.6501	0.5790	0.5061	0.2779	0.1673
	4	0.6523	0.5508	0.5193	0.3167	0.1523
	5	0.6927	0.5760	0.5489	0.2896	0.1477

Table 3. Grey relation degree between pending series matrix and standard fault matrix.

Sample	Normal	Slight-Worn	Medium-Worn	Broken-Teeth	Identification Result
1	0.8855	0.6501	0.7176	0.5410	Normal
2	0.9384	0.6775	0.7279	0.4927	Normal
3	0.8397	0.6154	0.7121	0.5656	Normal
4	0.9221	0.7196	0.7697	0.5748	Normal
5	0.8548	0.7678	0.8076	0.5329	Normal
6	0.6908	0.9083	0.8043	0.5865	Slight-worn
7	0.7188	0.9536	0.7704	0.5854	Slight-worn
8	0.7077	0.9338	0.7954	0.6209	Slight-worn

Table 3. Contd….

Sample	Normal	Slight-Worn	Medium-Worn	Broken-Teeth	Identification Result
9	0.6816	0.8753	0.7858	0.5107	Slight-worn
10	0.6867	0.9168	0.7522	0.6225	Slight-worn
11	0.6808	0.7450	0.8187	0.4540	Medium-worn
12	0.7004	0.6516	0.8518	0.3939	Medium-worn
13	0.7272	0.8063	0.9280	0.4611	Medium-worn
14	0.7457	0.7571	0.9138	0.4589	Medium-worn
15	0.7132	0.7716	0.9625	0.4686	Medium-worn
16	0.6104	0.6791	0.6678	0.7784	Broken-teeth
17	0.5145	0.5599	0.5170	0.8667	Broken-teeth
18	0.6017	0.6833	0.6131	0.8321	Broken-teeth
19	0.6409	0.5748	0.6593	0.8730	Broken-teeth
20	0.6457	0.6524	0.6399	0.8530	Broken-teeth

Table **2** gives five sample data of each data set selected randomly. Then we set the values of Table **1** as the standard fault set, and we recognize different gear fault mode by calculating the grey relation degree between the fault sample and standard fault pattern. Table **3** gives the final identification results. We can see that each fault pattern has been identified by the proposed method.

CONCLUSIONS

In this paper, a novel gearbox fault identification way is proposed by using mathematical morphological filter, EEMD, sample entropy and grey relation degree. Firstly, mathematical morphological filter is used to eliminate the noise interferences in original gearbox vibration signal. Secondly, EEMD is used to decompose the processed signal adaptively into a finite number of stationary intrinsic mode functions. Thirdly, the sample entropy of the first five IMFs containing the most dominant fault information is calculated and served as the fault feature. Finally, the grey relation degree between the fault sample and standard fault pattern is obtained as the evidence of fault identification. Practical examples verify that the proposed method is very useful for gearbox fault type diagnosis. It has a great application value in fault diagnosis.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This research is financially supported by Yunnan application fundamental research project fund (2013FB062) and Academic leader reserve talent project fund of Honghe University (2014HB0205). All the persons involved in the research projects are thanked for their help.

REFERENCES

[1] A. K. S. Jardine, D. Lin and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing*, vol. 20, pp. 1483-1510, September, 2009.

[2] C. D. Duan, Z. J. He and H. K. Jiang, "New method for weak fault feature extraction based on second wavelet transform and its application", *Chinese Journal of Mechanical Engineering*, vol. 17, pp. 543-547, April, 2004.

[3] Q. M. Zhang and H. J. Liu, "Processing of sampled power system data based on generalized morphological filtering", *Electric Power Automation Equipment*, vol. 26, pp. 45-48, October, 2006.

[4] M. Q. Xu, R. X. Wang, and J. Z. Hang, "De-Noising technique for vibration signals of rotating machine based on comb wavelet", *Journal of Vibration Engineering*, vol. 15, pp. 89-91, January, 2002.

[5] W. B. Zhang, X. J. Zhou and F. H. Mu, "Application of morphological filter in purification of rotor center's orbit", In: *Proceedings of the 2008 1st International Congress on Image and Signal Processing*, 2008, pp. 136-140.

[6] W. B. Zhang, X. J. Zhou and Y. Lin, "Application of morphological filter in pulse noise removing of vibration signal", In: *Proceedings of the 2008 1st International Congress on Image and Signal Processing*, 2008, pp. 132-135.

[7] D. E. Lake, J. S. Richman, M. P. Geriffin and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability", *Am J Physiology Regular Integrative and Comparative Physiology*, vol. 283, pp. 789-797, March, 2002.

[8] W. B. Zhang, L. Shen, J. S. Li, Q. Cai and H. J. Wang, "Morphological undecimated wavelet decomposition for fault feature extraction of rolling element bearing", In: *Proceedings of the 2009 2nd International Congress on Image and Signal Processing*, 2009, pp. 4252-4258.

[9] A. J. Hu, G. J. Tang, and L. S. An, "New method of Removing Pulsed Noises in Vibration Data", *Journal of Vibration and Shock*, vol. 25, pp. 126-127, January, 2006.

[10] Z. H. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method", *Advances in Adaptive Data Analysis*, vol. 1, pp. 1-41, January, 2009.

[11] L. Chen, G. S. Tang, Y. Y. Zi and F. Fan, "Improved EEMD applied to rotating machinery fault diagnosis", *Applied*

Mechanics and Materials, vol. 128-129, pp. 154-159, October, 2011.

[12] J. S. Lin and Q. Chen, "Application of the EEMD method to multiple faults diagnosis of gearbox", In: *2010 2nd International Conference on CD-ROM Purchase at Partner*, 2010, pp. 87-90.

[13] R. Alcaraz and J. J. Rieta, "Review on sample entropy applications for the non-invasive analysis of atria fibrillation electrocardiograms", *Biomedical Signal Processing and Control*, vol. 5, pp. 1-14, January, 2010.

[14] Y. G. Luo and C. Z. Chen, "Diagnostic analysis of grey network in rotary machinery trouble", *Compressor Blower & Fan Technology*, vol. 4, pp. 38-40, April, 2001.

[15] Y. B. Dong and X. L. Zhang, "Determination method for identification coefficient of grey relational grade and applying in mechanical faults diagnosis", *Equipment Manufacturing Technology*, vol. 3, pp. 121-125, March, 2008.

[16] Q. Zhao, J. Gao, T. F. Wu and L. Lu, "The grey theory and the preliminary probe into information acquisition technology", In: *Proceedings International Conference on Information Acquisition*, 2004, pp. 402-404.

Received: September 10, 2014 Revised: November 5, 2014 Accepted: November 5, 2014