A Data Mining based Full Ceramic Bearing Fault Diagnostic System using AE Sensors

David He, Ruoyu Li, Junda Zhu, and Mikhail Zade

Abstract - **Full ceramic bearings are considered the first step towards full ceramic, oil free engines in the future. No research on full ceramic bearing fault diagnostics using acoustic emission (AE) sensors has been reported. Unlike their steel counterparts, signal processing methods to extract effective AE fault characteristic features and fault diagnostic systems for full ceramic bearings have not been developed. In this paper, a data mining based full ceramic bearing diagnostic system using AE based condition indicators (CIs) is presented. The system utilizes a new signal processing method based on Hilbert Huang transform (HHT) to extract AE fault features for the computation of CIs. These CIs are used to build a data mining based fault classifier using a *k*-nearest neighbor (KNN) algorithm. Seeded fault tests on full ceramic bearing outer race, inner race, balls, and cage are conducted on a bearing diagnostic test rig and AE burst data are collected. The effectiveness of the developed fault diagnostic system is validated using real full ceramic bearing seeded fault test data.** [*]

Index Terms – Fault diagnosis, full ceramic bearings, data-mining .

I. INTRODUCTION

CERAMIC bearings are quickly replacing conventional steel ball bearings in various fields and applications because they exhibit a service life three times longer than that of steel bearings. There are two types of ceramic bearings: hybrid ceramic bearings and full ceramic bearings. Hybrid ceramic bearings have steel races and ceramic balls while full ceramic bearings have both ceramic balls and races. Different types of ceramics are used in ceramic bearings. Silicon nitride (Si3N4) and Zirconia (ZrO2) are perhaps the most common ceramics used in ceramic bearings. However there are many other ceramics that would work well in bearing applications (Niizeki, 2000).

Very few studies have been conducted on ceramic bearing fault diagnosis and remaining useful life prognosis of hybrid ceramic ball bearings. Past research has focused on studying the characteristics and applications of ceramic bearings (Wereszczak *et al.*, 2006; Takebayashi, 2001; Arulampalam and Ristic, 2000; Chao *et al.*, 1995; Ohta and Kobayashi, 1995; Rhoads and Bashyam, 1994; Ebert, 1990). Recently, a number of papers have reported research on hybrid ceramic

D. He, R. Li, J. Zhu, and M. Zade are with the Intelligent Systems Modeling & Development Lab in the Department of Mechanical and Industrial Engineering at The University of Illinois at Chicago, Chicago, IL 60607, USA. Email: davidhe@uic.edu (D. He), rli8@uic.edu (R. Li), jzhu26@uic.edu (J. Zhu), mzade2@uic.edu (M. Zade).

bearing fault diagnosis and prognosis. Dempsey *et al.* (2004, 2005) summarized the currently known failure modes of hybrid ceramic bearings. In their research, they used both magnetic and non-magnetic sensors to detect the silicon nitride debris. A hybrid bearing test rig has been developed by National Aeronautics and Space Administration (NASA) at Glenn Research Center in order to evaluate the performance of the sensors and algorithms developed in predicting failures of rolling element bearings for aeronautic and space applications. Different diagnostic tools, both oil based and vibration based systems, were used to indicate bearing failures. The vibration data were recorded and analyzed in time domain and frequency domain and envelope analysis techniques developed to indicate the health condition of bearings in real-time. In the meanwhile, several different oil debris sensors were installed to get the information of both metallic and non-metallic debris particles. Using the magnetic properties of the oil debris to detect damage is not enough since the ceramic rolling elements of hybrid bearings have no metallic properties.

Byington *et al.* (2006) presented a feature extraction and analysis driven system: ImpactEnergy. This system recorded high frequency vibration/acoustic emission data and combined advanced diagnostic features derived from waveform analysis, high-frequency enveloping, and more traditional time domain processing like root mean square (RMS) and kurtosis with classification techniques to provide bearing health information. However, the fault detection and diagnostic algorithms were the focus of the paper. The prediction algorithm of remaining useful life of hybrid ceramic bearings was not reported and verified in the paper. He *et al.* (2010) presented a methodology for hybrid ceramic bearing prognostics using particle filtering. The methodology was validated using real hybrid ceramic bearing run to failure test data collected on a bearing prognostics test rig. The collected data were used to test the prognostic algorithm. The results have shown that the presented prognostics methodology is effective for predicting the remaining useful life of the tested bearings. Ma and He (2010) presented an integrated prognostics methodology and illustrated the method using real and simulated run-to-failure test data on hybrid ceramic bearings. The results show that when grease debris and vibration analysis are integrated within one framework, the trending of the fault features can be improved significantly. The prediction uncertainty can be reduced. This integrated methodology could enhance the machine condition monitoring performance and make remaining useful life prediction more reliable.

Recently, acoustic emission (AE) based techniques have attracted researchers' attention to machine health monitoring and fault diagnostics (Eftekharnejad and Mba, 2009; Hamzah and Mba, 2009; Samuel and Pines, 2005; Toutountzakis *et al.*, 2005; Choudhury and Tandon, 2000). Tandon and Nakra (1990) used AE counts and peak amplitudes of continuous AE signals as the fault features for steel bearing outer race fault diagnostics. Tan *et al.* (2007) compared fault diagnostic and prognostic capabilities of AE and vibration based techniques, and spectrometric oil analysis (SOA). They used RMS values as the features for both the AE and vibration signals. Their results showed that the AE based techniques were more sensitive in detecting and monitoring pitting than either the vibration or SOA techniques. In (Mba, 2008), the duration of the AE bursts were used to estimate the defect size of the steel roller bearing outer race.

Up to today, no research on full ceramic bearing fault diagnosis using AE sensors has been reported. Unlike their steel counterparts, effective signal processing methods to extract AE fault characteristic features and fault diagnostic systems for full ceramic bearings have not been developed. The applications of full ceramic bearings can be limited due to the lack of fault diagnostic technology, especially in the fields of aeronautics and aerospace applications. Reliable diagnostics tools and algorithms should be developed.

In general, there are two popular approaches for bearing fault diagnostics. One is physics based or model based approach. This approach establishes a mathematical model of the bearings and then by monitoring the residual signals between the model and the collected signals it classifies the bearing faults. Some model based bearing fault detection and diagnosis applications have been reported in (Jalan and Mohanty, 2009; Blodt *et al.*, 2008; Stack *et al.*, 2004; Loparo *et al.*, 2000). However for a brand new or complex bearing system, it is infeasible to obtain an accurate model. Therefore, the model-based approach is limited. Another drawback of

the model-based approach is that different models need to be established for different applications. Another popular approach is data mining based method. Some data mining based bearing fault detection and diagnosis applications have been reported in (Abbasion *et al.*, 2007; Samanta *et al.*, 2003; Jack and Nandi, 2002; Li, *et al.*, 2000). Unlike model-based approach, the data mining based approach does not necessarily need to understand the physics of the monitored applications. Thus the complexity of designing the methodology is greatly reduced, especially for the new and complex systems. Data mining based methods can be easily applied to complex systems and adapted to applications with different system parameters. Among the data mining approaches, neural networks based methods as an effective learning tool have been reported (Ho *et al.*, 2010; Sun and Kabán, 2010; Talebi *et al.*, 2009).

In this paper, a data mining based full ceramic bearing diagnostic system using AE based condition indicators (CIs) is presented. The system utilizes a new signal processing method based on Hilbert Huang transform (HHT) to extract AE fault features for the computation of CIs. These CIs are used to build a data mining based fault classifier using a *k*-nearest neighbor (KNN) algorithm. Seeded fault tests on full ceramic bearing outer race, inner race, balls, and cage are conducted on a bearing diagnostic test rig and AE burst data are collected. The effectiveness of the developed fault diagnostic system is validated using real full ceramic bearing seeded fault test data.

II. DATA MINING BASED FAULT DIAGNOSTIC SYSTEM USING AE SENSORS

The framework of the data mining based full ceramic bearing fault diagnostic system using AE sensors is provided in Figure 1.
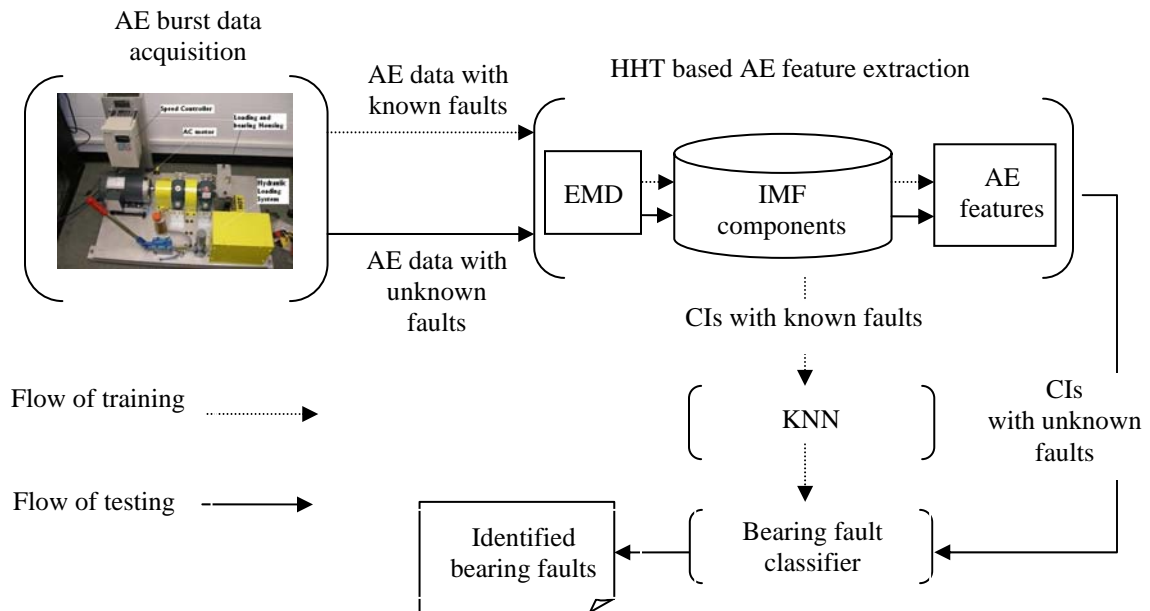

Fig. 1 Data mining based full ceramic bearing fault diagnostic system using AE sensors.

As shown in Figure 1, AE bursts from both healthy and faulty bearings are collected and then decomposed into intrinsic mode functions (IMF) components by a technique called empirical mode decomposition (EMD) method. As pointed out by Everson and Cheraghi (1999), the AE signals could be generated as two types: burst type and continuous type. Burst type signals contain pulses over a period of time above background signals and can be separated with limited overlap. Continuous type signals do not have any significant resolutions between individual pulses. The AE signals generated by a rotating component with localized faults such as cracks belong to the burst type signals (Dolinek and Kopa, 1999). One major advantage of using burst type AE signals for bearing fault diagnosis is that shorter sampling periods will be needed. Therefore fewer storage spaces and less computational burden will be added to the real monitoring systems. In general, bearing faults such as those damages on inner race, outer race, and balls can be treated as localized faults. The localized faults not only affect the transmission accuracy but also cause catastrophic failure of the transmission system. In this paper, the diagnosis of bearing localized faults on ceramic inner race, outer race and balls are emphasized.

After the EMD decomposition, the AE based CIs are then computed from the obtained IMF components. These AE based CIs are used to train a data mining based bearing fault classifier using a k-nearest neighbor (KNN) algorithm.

A. HHT based AE Feature Extraction and CI Development

HHT was first proposed by Huang (Huang *et al*., 1998) and has been recently applied to rotational machinery fault detection (Liu *et al*., 2006; Yan and Gao, 2006) and short circuit fault detection in a permanent-magnet synchronous motor (Rosero *et al*., 2009). It has been proven to be an effective method in analyzing non-stationary signals for rotational machinery fault detection. When HHT is applied to analyze the signals, the signals are first decomposed into several IMF components by EMD method. The steps of the EMD method are provided as follows:

Step 1. Find the local maxima and local minima of the signals.

Step 2. Construct the lower and upper envelopes of the signals by the cubic spline respectively based on the local maxima and local minima.

Step 3. Calculate the mean values $m(t)$ by averaging the lower envelope and the upper envelope.

Step 4. Subtract the mean values from the original signals to produce the IMF candidate component $h_1(t) = f(t) - m(t)$. If it is the true intrinsic mode function, go to the next step and the IMF component $C_i(t) = h_m(t)$ is saved. If it is not the intrinsic mode function, go back to Step 1. The stop condition for the iteration is given by Eq. (1):

$$\sum_{t=0}^{T} \frac{[h_{m-1}(t) - h_m(t)]^2}{h_{m-1}^2(t)} \leq SD \qquad (1)$$

where: $h_{m-1}(t)$ and $h_m(t)$ denote the IMF candidates of the m-1 and m iterations, respectively and usually SD is set between 0.2 and 0.3.

Step 5. Calculate the residual component by subtracting IMF component obtained in Step 4 from the original signals $res_i(t) = f(t) - C_i(t)$. This residual component is treated as new data and is subjected to the same processes described above to calculate the next IMF component.

Step 6. Repeat Steps 1-5 until the final residual component becomes a monotonic function and no more IMF component can be extracted or the envelopes becomes smaller than a pre-determined value.

Through the EMD, the original signals $f(t)$ can be decomposed into N IMF components: C_1, ..., C_i, ..., C_N. From these IMF components, AE features can be extracted. In this paper, three AE features are investigated: *rms*, *kurtosis*, and *peak value*. Let $x = (x_1, ..., x_i, ..., x_n)$ be a set of data, $\bar{x}$ the average, and s the standard deviation. Then from dataset x, the three AE features can be computed as:

$$rms = \sqrt{\frac{\sum_i^n x_i^2}{n}} \qquad (2)$$

$$kurtosis = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)} \qquad (3)$$

$$peak\ value = \max_i \{x_i, i = 1, ..., n\} - \min_i \{x_i, i = 1, ..., n\} \qquad (4)$$

The AE feature *rms* captures the energy level of the AE bursts. A faulty bearing normally has a higher energy level than a normal bearing under the same operating conditions. The AE feature *kurtosis* indicates the flatness or the spikiness of the AE bursts. Its value is very low for normal condition of the bearing and high for faulty condition of the bearing due to the spiky nature of the burst. The AE feature *peak value* captures the relative flatness or the spikiness of the AE bursts.

Once the AE features are extracted from the IMF components of the AE bursts, CIs can be generated from these AE features. CIs have been widely used in machinery health monitoring and fault diagnosis applications. A large number of vibration based CIs are used in health and usage monitoring systems (HUMS) currently installed in military and civil service helicopters for health monitoring and fault diagnosis of flight critical components such as bearings, gears, and engines. For example, some vibration based CIs for bearing health monitoring and fault diagnosis are defined using the fundamental bearing fault frequencies: outer race ball-pass

frequency (BPFO), inner race ball-pass frequency (BPFI), ball spin frequency (BSF), and cage fault frequency (CFF) or fundamental train frequency (FTF). These CIs are used to detect bearing faults in bearing health monitoring and fault diagnosis applications.

Unlike vibration analysis, AE based bearing fault features and CIs have not been well established yet, especially for full ceramic bearings. Since the collection of each AE burst data is triggered by AE impulses, the bearing fundamental frequencies used in vibration analysis cannot be computed using AE burst data. In this paper, simple and effective CIs computed from AE burst data are developed. Assume that the AE features can be arbitrarily divided into finite number of groups of size M. Let f_i be the value of AE feature i. Then the following CIs can be computed for each AE features as follows:

$$group\ average\ of\ AE\ feature\ i = f_i = \frac{\sum_{i=1}^{M} f_i}{M} \tag{5}$$

$$group\ stdev\ of\ AE\ feature\ i = \sqrt{\frac{\sum_{i=1}^{M}\left(f_i - \overline{f}_i\right)^2}{M-1}} \tag{6}$$

The CI in Eq. (5) is basically computed as the group average of an AE feature. It will indicate the magnitude of the AE impulses caused by the bearing faults. The CI in Eq. (6) is basically computed as the group standard deviation of an AE feature. It will indicate the randomness of the AE impulses caused by the bearing faults. It is expected that this CI should have lower values for healthy bearing and bearing with faults on non-rotating components such as outer race. Once the CIs of the AE bursts for each type of bearing faults are generated, they will be used to train the bearing fault classifier using a KNN algorithm.

B. The KNN Algorithm

KNN is a supervised learning algorithm. KNN assumes all observations correspond to points in a p-dimensional space. The nearest neighbors of an observation are defined in terms of the standard Euclidean distance. An observation is classified by a majority vote of its neighbors, with the observation being assigned the class most common amongst its k nearest neighbors. Consider a database consisting of a total of n observations ($\mathbf{x}_i$; y_i), for $i = 1, 2, …, n$, where $\mathbf{x}_i$ could be any point in the p-dimensional Euclidean space, $\Re^p$, denoted as $\mathbf{x}_i = \{x_{i1}, x_{i2}, ..., x_{ip}\}$ and y_i is an outcome from m class $\omega = \{\omega_1, \omega_2, ..., \omega_m\}$. The database is called the training set for the KNN algorithm. Given any two observations, $\mathbf{x}_i$ and $\mathbf{x}_j$, let $s(\mathbf{x}_i, \mathbf{x}_j)$ be a measure of their similarity based on the p variates and be derived as:

$$s\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sqrt{\sum_{k=1}^{p}\left(x_{ik} - x_{jk}\right)^2} \tag{7}$$

To classify the response for a new observation $\mathbf{x}_u$ with the KNN algorithm, one first identifies k observations in the training set that are most similar to $\mathbf{x}_u$. They form the set of k-nearest neighbors of $\mathbf{x}_u$, denoted by $N\left(\mathbf{x}_u, k\right)$. These similarities can be ordered. Denote the ordered similarities with $s_{(i)}$, i.e., $s_{(1)} \geq s_{(2)} \geq ... \geq s_{(n)}$. In other words, if $s_j = s_{(k)}$, it means $\mathbf{x}_j$ is the kth most similar observation in the training set to $\mathbf{x}_u$. The set of the k-nearest neighbors of $\mathbf{x}_u$, $N\left(\mathbf{x}_u, k\right)$, can then be defined as all observations whose similarities to $\mathbf{x}_u$ are at least $s_{(k)}$, i.e., $N\left(\mathbf{x}_u, k\right) = \left\{x_i : s_i \geq s_{(k)}\right\}$. The KNN algorithm then estimates the probability that $y_u = \omega_i$ by the average responses of these k-nearest neighbors and classifies the response to be ω_i if the estimated probability exceeds a certain threshold c:

$$\hat{P} = \frac{\sum_{x_i \in N(\mathbf{x}_u, K)} y_i}{\left|N\left(\mathbf{x}_u, K\right)\right|} \tag{8}$$

where: $\left|N\left(\mathbf{x}_u, k\right)\right|$ is the number of items contained in the set $N\left(\mathbf{x}_u, k\right)$. This is usually equal to k exactly, but may exceed k depending on how ties are treated. The response is then predicted to be ω_i if $\hat{P} \geq c$, where c is a pre-specified threshold parameter.

III. EXPERIMENTAL SETUP

In order to validate the AE based full ceramic bearing fault diagnosis method, bearing seeded fault tests were conducted on a specially designed bearing test rig (see Figure 2) and AE burst data were collected.
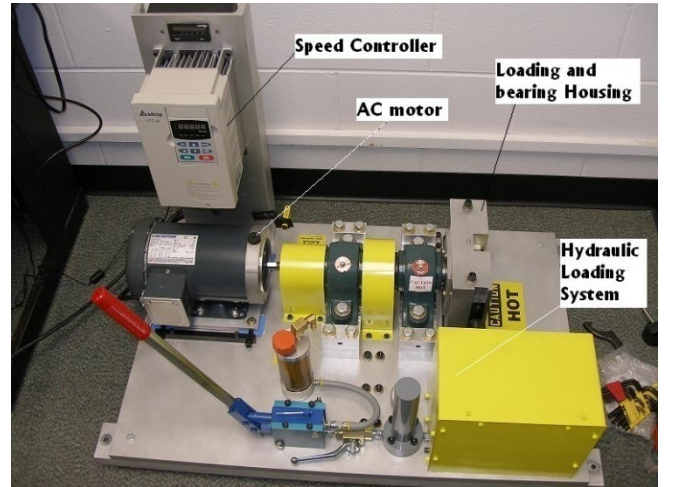

Fig. 2 The bearing test rig.

The bearings used during the testing were the 6205 full ceramic bearings purchased from Boca Bearings. The bearing balls, outer race, and inner race are made of Silicon Nitride (Si3N4). The bearing cage (retainer) and seals are made of Teflon. The lube for the full ceramic bearing is LD lube dry. The dimension of the bearing is shown in Figure 3.
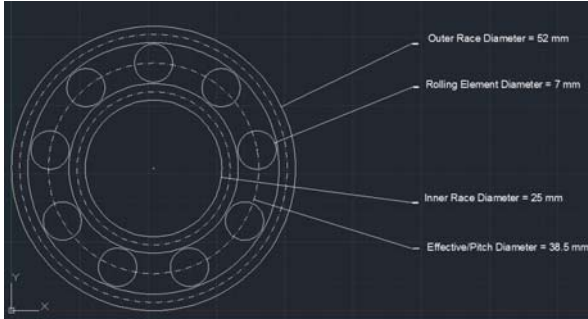

Fig. 3 The dimension of the 6205 full ceramic bearing.

To simulate the early faults developed on a bearing, 4 different types of full ceramic bearing faults were created: damaged inner race surface, damaged outer race surface, damaged ball, and broken cage (see Figure 4).

inner race fault   outer race fault

ball fault   cage fault

Fig. 4 The four bearing seeded faults.

The surface damages on the bearing inner race and outer race were created by scratching the race surface using a grinding tool with a diamond tip. The diameter of the damaged surface area was about a quarter of the ball diameter. The surface damage on one bearing ball was created by scratching the ball surface using the grinding tool. Roughly 10% of ball surface were scratched. The broken cage was created by cutting the Teflon retainer using a sharp knife.

Two wide band (WD) type AE sensors and a 2-channel data acquisition card with 18-bit resolution and a maximum

sampling rate up to 40 MHz were used to collect the AE burst data. The AE sensors were attached to the bearing housing by instant glues. The locations of the AE sensors on the bearing housing are shown in Figure 5.

Fig. 5 The locations of the AE sensors.

During the testing, the speed of the motor shaft was controlled at 10Hz (600 rpm) and the AE burst data were collected at a sampling rate of 5 MHz for the 5 full ceramic bearings: healthy bearing, bearing with inner race fault, bearing with outer race fault, bearing with ball fault, and bearing with cage fault. For each bearing, a total of 400 AE bursts were collected in one data acquisition. In each data acquisition, the bearing was run for less than 2 minutes. The time to record a total of 400 AE bursts was less than a minute. It is assumed that the short data acquisition duration did not permit fault propagation during the testing.

To keep a consistent testing condition, only one bearing house was used for all five bearings. Every time when a bearing test was completed, the tested bearing was taken out from the bearing house and the next bearing was put into the same bearing house. As the position of bearing house in the bearing test rig and the AE sensors attached to the bearing house did not change each time when the tested bearing was replaced, the variability of the testing conditions was minimized

IV. ANALYSIS RESULTS

For each AE burst data, the EMD method was applied to decompose it into a number of IMF components. An example of AE bursts and the first 4 IMF components of the AE burst are provided in Figure 6.

As one can see from Figure 6 that for a typical AE burst collected during the full ceramic bearing seeded fault testing, the first three IMF components contained most of the energy in the original AE burst. Therefore, only the first three IMF components of the AE bursts were used to extract the AE features. To extract the AE features, the three IMF components were summed and then the three AE features were extracted from the summed IMF components: *rms*, *peak*

value, and *kurtosis*. For each AE feature, the first 10 values were grouped as the first data group, the second 10 values the second data group, and this process of putting the next 10 values into a data group continued until 40 data groups were generated. For each data group, the group average and standard deviation were computed. Treating each group average and standard deviation as CIs, then a total of 6 CIs were generated.

(a) AE burst

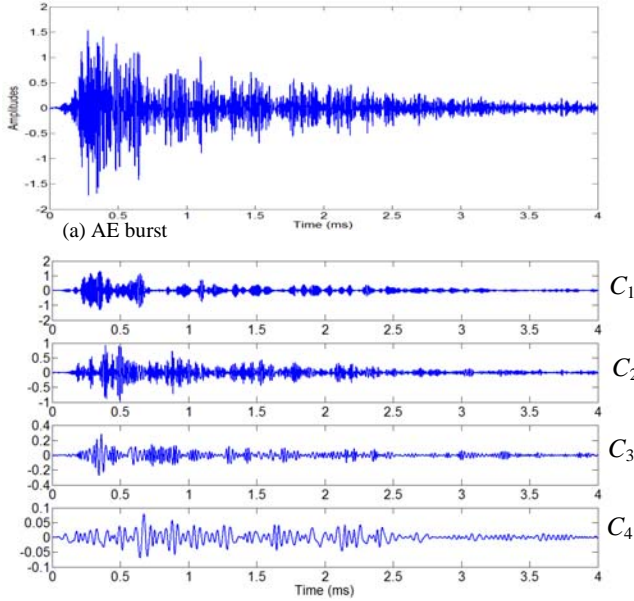(b) The first 4 IMF components of the AE burst in (a)

Fig. 6 An AE burst and its first four IMF components.

The group averages and the group standard deviations of the AE feature *rms* of the five tested bearings are plotted in Figure 7 and Figure 8, respectively.
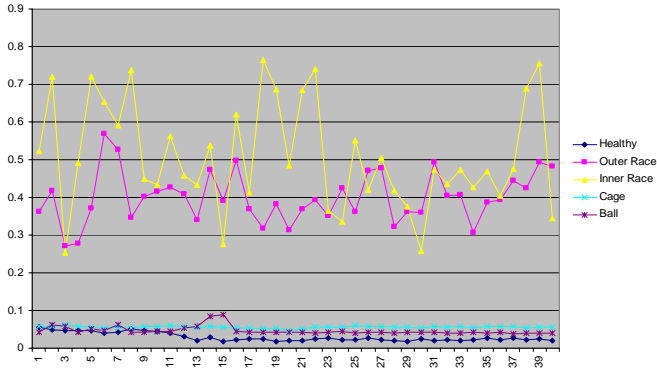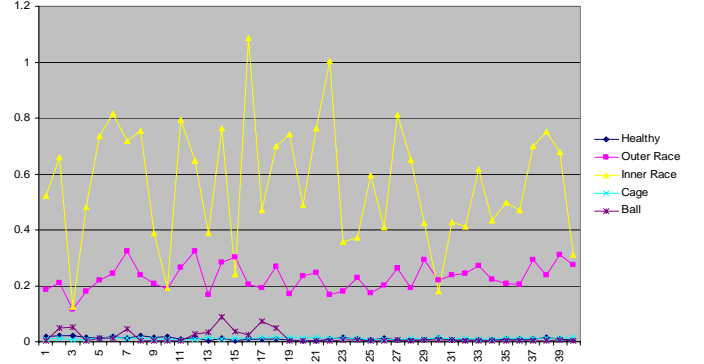

Fig. 7 The group averages of *rms*.

Fig. 8 The group standard deviation of *rms*.

From Figure 7, one can see that the AE burst energy level of inner race and outer race faults measured by the group average of *rms* is much higher than that of ball and cage faults. This is because the impacts generated by the ball hitting on the damaged surface of the inner and outer race are much stronger. Therefore, by using the group average of *rms* alone, one can separate the race faults from ball and cage faults. Since the AE burst energy levels of the inner race fault and outer race fault are overlapped, the group average of *rms* cannot separate the outer race fault from the inner race fault. However, from Figure 8, one can see that using the group standard deviation of *rms* one could roughly separate the inner race fault from the outer race fault. This is because the group standard deviation of *rms* measures the randomness of the AE impulses. Since the AE impulses of the inner race fault were generated by the ball running against the rotating inner race rather than the non-rotating outer race, they show more randomness than the AE impulses of the outer race fault.

The group averages and the group standard deviations of the AE feature *kurtosis* of the five tested bearings are plotted in Figure 9 and Figure 10, respectively.
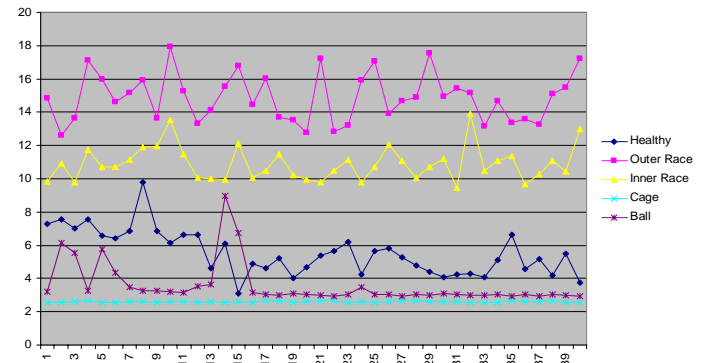

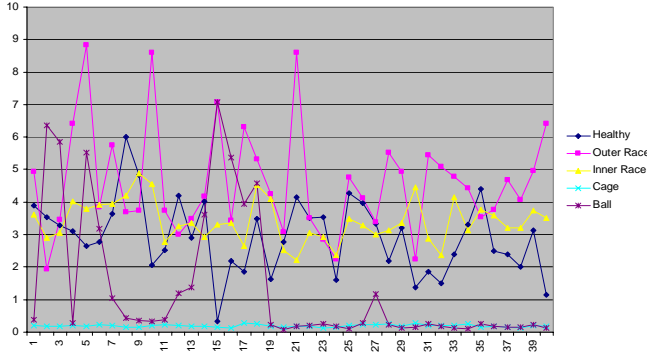Fig. 9 The group averages of *kurtosis*.

6

Fig. 10 The group standard deviations of *kurtosis*.

From Figure 9, one can see that except for data points #14 and #15 at which the CI values of the ball fault are higher than those of the healthy bearing, all the five bearings can be clearly separated by the group average of *kurtosis*. Hence, by using the group average of *kurtosis* alone, one could roughly separate the bearing faults. In theory, the kurtosis of a healthy bearing should have a lower value than a faulty bearing as the kurtosis measures the spikiness of the signals. However, note that in Figure 9, the group averages of *kurtosis* of the healthy bearing are higher than those of the ball fault and cage fault. This might be caused by the fact that the seeded faults created on the surface of the ball and the cage were not large enough and the resulting pulses generated by these faults were easily buried in the background noises and overlapped with regular AE bursts. From Figure 10, one cannot see a good separation of the bearing faults using the group standard deviation of *kurtosis*.

The group averages and the group standard deviations of the AE feature *peak value* of the five tested bearings are plotted in Figure 11 and Figure 12, respectively.
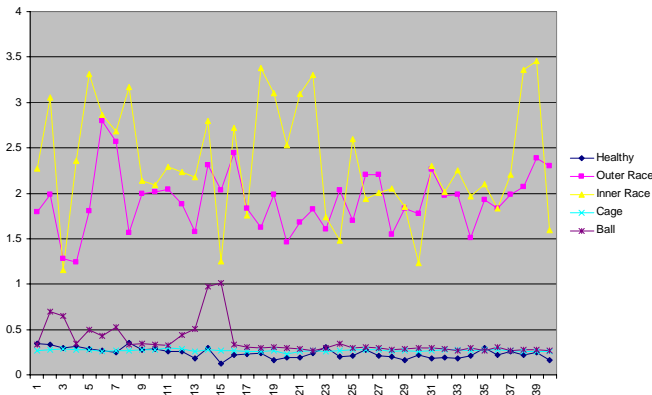

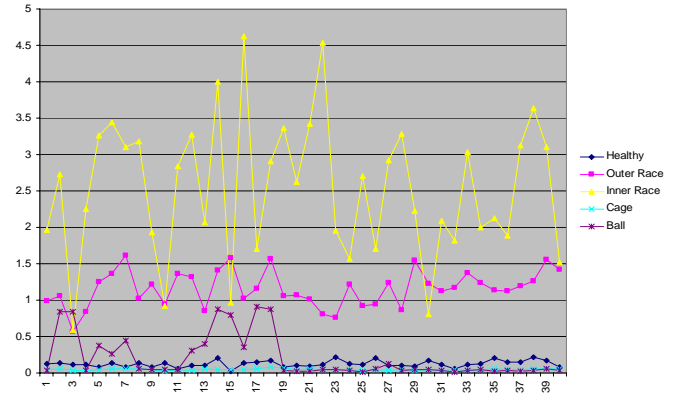Fig. 11 The group averages of *peak value*.

Fig. 12 The group standard deviations of *peak value*.

From Figure 11, one can see that the AE burst peak level of inner race and outer race faults measured by the group average of *peak value* is much higher than that of ball and cage faults. This is because the AE burst peaks generated by the ball hitting on the damaged surface of the inner and outer races are much higher. Therefore, by using the group average of *peak value* alone, one can separate the race faults from ball and cage faults. However, the group average of *peak value* cannot separate the outer race fault from the inner race fault. From Figure 12, one can see that by using the group standard deviation of *peak value* one could roughly separate the inner race fault from the outer race fault.

In summary, the best CI seems to be the group average of *kurtosis* as it shows relatively good separation among different bearings in comparison to other CIs. For both *rms* and *peak value*, their CIs seems to be able to separate race faults from non-race faults but unable to separate bearings that don't have the race faults: healthy bearing, bearing with ball fault, and bearing with cage fault.

As reported in the data mining based fault diagnostic papers by Kim and Parlos (2002), Wu and Chow (2004), and Sreejith *et al.* (2008), a good practice of splitting the data for training and validation is that the percentage of data used for training should be higher than 50%. Therefore, in this paper, all the AE CI data generated were split into two groups: 60% of the data were used for training purpose to build up the fault classifier and 40% of the data were used for validation to test the performance of the AE based CIs. The fault diagnosis results using KNN for different CIs and CI combinations are provided in Table I and Table II, respectively. Note that in both Table I and Table II, the values of *k* are the best *k* values with the minimum % error determined by the KNN algorithm at the training stage. The % error is defined as the percentage of data points of a fault type that were misclassified.

TABLE I
THE KNN FAULT DIAGNOSIS RESULTS USING SINGLE CI ONLY

CI: *rms*-average ($k = 6$)			CI: *rms*-standard deviation ($k = 8$)		
Type of Fault	% Error for Training	% Error for Validation	Type of Fault	% Error for Training	% Error for Validation
Healthy	13.04	11.76	Healthy	21.74	11.76
Outer race	4.00	6.67	Outer race	12.00	20.00
Inner race	12.50	18.75	Inner race	66.67	50.00
Ball	19.23	21.43	Ball	7.69	21.43
Cage	36.36	38.89	Cage	0.00	0.00
Overall	16.67	20.00	Overall	21.67	20.00
CI: *kurtosis*-average ($k = 5$)			**CI: *kurtosis*-standard deviation ($k = 4$)**		
Type of Fault	% Error for Training	% Error for Validation	Type of Fault	% Error for Training	% Error for Validation
Healthy	8.70	17.65	Healthy	21.74	47.06
Outer race	0.00	0.00	Outer race	36.00	20.00
Inner race	4.17	18.75	Inner race	20.83	50.00
Ball	0.00	21.43	Ball	38.46	57.14
Cage	0.00	0.00	Cage	59.09	72.22
Overall	2.50	11.25	Overall	35.00	50.00
CI: *peak value*-average ($k = 5$)			**CI: *peak value*-standard deviation ($k = 4$)**		
Type of Fault	% Error for Training	% Error for Validation	Type of Fault	% Error for Training	% Error for Validation
Healthy	8.70	35.29	Healthy	17.39	29.41
Outer race	4.00	26.67	Outer race	16.00	33.33
Inner race	41.67	31.25	Inner race	4.17	12.50
Ball	26.92	35.71	Ball	11.54	21.43
Cage	13.64	27.78	Cage	13.64	16.67
Overall	19.17	31.25	Overall	12.50	22.50

The results shown in Table I are consistent with the observations of the CI plots shown in Figure 7 through Figure 12. The group average of *kurtosis* has the lowest classification error among all the individual CIs tested. However, none of the individual CI tested has achieved a classification accuracy rate over 90%.

The results in Table II have shown that the classification accuracy can be significantly improved by combining individual CIs. For example, by combining group average of *rms* with group standard deviation of *rms*, the classification error was reduced significantly from 20% to 6.25%. In the case of combining all the CIs, the error was reduced to 7.50%. Overall, by combining all individual CIs, a classification accuracy rate of over 92% can be achieved.

In this paper, the performance of the KNN based fault classifier was compared with a neural network based classifier. For the comparison purpose, the neural network based classifier was used to classify the bearing faults using the same CIs as the KNN algorithm. The neural network used for comparison was a backpropagation neural network with one hidden layer. To train the neural network, 1000 iterations were used. The fault diagnosis results using the neural network for the same CI combinations shown in Table II are provided in Table III. Note that in Table III, the values of *Node* are the best number of nodes of the hidden layer with the minimum % error determined by the neural network algorithm at the training stage. The % error is defined as the percentage of data points of a fault type that were misclassified. From the results shown in Table III, a similar conclusion to that by the KNN based algorithm can be drawn. The classification accuracy can be significantly improved by combining all individual CIs. By comparing with the results shown in Table II and Table III, one could see that the performance of the KNN based algorithm is comparable with that of the neural network based algorithm. Over 92% accuracy rate was achieved by both algorithms when all the CIs were combined.

TABLE II
THE KNN FAULT DIAGNOSIS RESULTS USING COMBINATION OF CIS

CIs: *rms*-average + *rms*-standard deviation ($k = 4$)		
Type of Fault	% Error for Training	% of Error for Validation
Healthy	8.70	5.88
Outer race	4.00	0.00
Inner race	4.17	12.50
Ball	3.85	14.29
Cage	9.09	0.00
Overall	5.83	6.25
CIs: *kurtosis*-average + *kurtosis*-standard deviation ($k = 3$)		
Type of Fault	% Error for Training	% of Error for Validation
Healthy	4.35	17.65
Outer race	0.00	0.00
Inner race	0.00	18.75
Ball	0.00	21.43
Cage	4.55	0.00
Overall	1.67	11.25
CIs: *peak value*-average + *peak value*-standard deviation ($k = 3$)		
Type of Fault	% Error for Training	% of Error for Validation
Healthy	21.74	35.29
Outer race	0.00	13.33
Inner race	0.00	12.50
Ball	3.85	7.14
Cage	9.09	11.11
Overall	6.67	16.25
CIs: *rms*-average + *rms*-standard deviation+ *kurtosis*-average + *kurtosis*-standard deviation + *peak value*-average + *peak value*-standard deviation ($k = 1$)		
Type of Fault	% Error for Training	% of Error for Validation
Healthy	0.00	11.76
Outer race	0.00	0.00
Inner race	0.00	6.25
Ball	0.00	14.29
Cage	0.00	5.56
Overall	0.00	7.50

TABLE III
THE ANN FAULT DIAGNOSIS RESULTS USING COMBINATION OF CIS

CIs: *rms*-average + *rms*-standard deviation (*Node* = 2)	
Type of Fault	% of Error for Validation
Healthy	31.25
Outer race	0.00
Inner race	14.29
Ball	88.24
Cage	0.00
Overall	27.50
CIs: *kurtosis*-average + *kurtosis*-standard deviation (*Node* = 10)	
Type of Fault	% of Error for Validation
Healthy	18.75
Outer race	0.00
Inner race	21.43
Ball	29.41
Cage	0.00
Overall	13.75
CIs: *peak value*-average + *peak value*-standard deviation (*Node* = 6)	
Type of Fault	% of Error for Validation
Healthy	87.50
Outer race	0.00
Inner race	14.29
Ball	70.59
Cage	0.00
Overall	35.00
CIs: *rms*-average + *rms*-standard deviation+ *kurtosis*-average + *kurtosis*-standard deviation + *peak value*-average + *peak value*-standard deviation (*Node* = 25)	
Type of Fault	% of Error for Validation
Healthy	6.25
Outer race	0.00
Inner race	21.43
Ball	5.88
Cage	0.00
Overall	6.25

V. CONCLUSIONS

Full ceramic bearings are considered the first step towards full ceramic, oil free engines in the future. No research on full ceramic bearing fault diagnostics using acoustic emission sensors has been reported. In this paper, a data mining based full ceramic bearing diagnostic system using AE based condition indicators was presented. Seeded fault tests on bearing outer race, inner race, ceramic balls, and cage were conducted and AE burst data were collected. Three AE features were extracted from the IMF components obtained using the EMD method. The three AE features were: *rms*, *kurtosis*, and *peak value*. Each AE feature was further processed using data groups to obtain the group averages and group standard deviations. The obtained group averages and standard deviations were treated as CIs to train a data mining based bearing fault classifier using a KNN algorithm. 60% of the CIs were used for training the fault classifier and 40% for validation to test the performance of the AE based CIs. The results of the validation have shown: (1) Among all the individual CIs tested, the group average of *kurtosis* has the lowest classification % error. However, none of the individual CIs tested has achieved a classification accuracy rate over 90%. (2) The classification accuracy can be significantly improved by combining individual CIs. The best combination of CIs is the combination of the group average of *rms* and group standard deviation of *rms*. The classification accuracy of this CI combination is over 92%. The performance of the KNN based fault classifier was also

compared with that of a backpropagation neural network based fault classifier. The comparison results showed that the performance of the KNN based algorithm is comparable with that of the neural network based algorithm. Over 92% accuracy rate was achieved by both algorithms when all the CIs were combined.

REFERENCES

[1] S. Abbasion, A. Rafsanjani, A. Farshidianfar, and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine", *Mechanical Systems and Signal Processing*, Vol. 21, pp. 2933 - 2945, 2007.

[2] S. Arulampalam and B. Ristic, "Comparison of the particle filter with range parameterized and modified polar EKFs for angle-only tracking", *Proc. SPIE*, Vol. 4048, pp. 288–299, 2000.

[3] M. Blodt, P. Granjon, B. Raison, and G. Rostaing, "Models for bearing damage detection in induction motors using stator current monitoring", *IEEE Transactions on Industrial Electronics*, Vol. 55, pp. 1813 - 1822, 2008.

[4] S. C. Byington, R. Orsagh, K. Pattada, J. Sheldon, M. DeChristopher, S. Amin, and J. Hines, "Recent Case Studies in Bearing Fault Detection and Prognosis", *IEEE Aerospace Conference*, Big Sky, MT, 2006.

[5] L. Y. Chao, D. K. Shetty, J. H. Adair, and J. J. Mecholsky, "Development of Silicon Nitride for Rolling Contact Bearing Applications: A Review", *Journal of Materials Education*, Vol. 17, No. 5-6, pp. 245–303, 1995.

[6] A. Choudhury and N. Tandon, "Application of acoustic emission technique for the detection of defects in rolling element bearings," *Tribology International,* vol. 33, pp. 39-45, 2000.

[7] P. J. Dempsey, J. M. Certo, and W. Morales, "Current Status of Hybrid bearing Damage Detection", *NASA/TM*—2004-212882, NASA, Cleveland, OH, 2004.

[8] P. J. Dempsey, M. C. Joseph, F. H. Robert, and D. Florin, "Hybrid Bearing Prognostic Test Rig", *NASA/TM*—2005-213597, NASA, Cleveland, OH, 2005.

[9] S. Dolinek and J. Kopa, "Acoustic emission signals for tool wear identification," *Wear*, Vol. 225, pp. 295-303, 1999.

[10] F. J. Ebert, "Performance of Silicon Nitride Components in Aerospace Bearing Applications", *ASME Publication*, No. 90-GT-166, 1990.

[11] B. Eftekharnejad and D. Mba, "Seeded fault detection on helical gears with acoustic emission," *Applied Acoustics,* Vol. 70, pp. 547-555, 2009.

[12] C. Everson and S. H. Cheraghi, "The application of acoustic emission for precision drilling process monitoring," *International Journal of Machine Tools and Manufacture*, Vol. 39, pp. 371-387, 1999.

[13] R. R. Hamzah and D. Mba, "The influence of operating condition on acoustic emission (AE) generation during meshing of helical and spur gear," *Tribology International,* vol. 42, pp. 3-14, 2009.

[14] D. He, A. Panyala, R. Li, and J. Ma, "The Use of Vibration Signals for Ceramic Bearing Prognostics", *Proceedings of the 2010 Industrial Engineering Research Conference*, Cancun, Mexico, June 5 – 9, 2010.

[15] K.I.-J. Ho, C. Leung, and J. Sum, "Convergence and objective functions of some fault/noise-injection based online learning algorithms for RBF networks", *IEEE Transactions on Neural Networks*, Vol. 21, No. 6, pp. 938 – 947, 2010.

[16] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A*, Vol. 454, pp. 903-995, 1998.

[17] L. Jack and A. Nandi, "Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms", *Mechanical Systems and Signal Processing*, Vol. 16, pp. 373 - 390, 2002.

[18] A. K. Jalan and A. Mohanty, "Model based fault diagnosis of a rotor-bearing system for misalignment and unbalance under steady-state condition", *Journal of Sound and Vibration*, Vol. 327, pp. 604 – 622, 2009.

[19] K. Kim and A. G. Parlos, "Induction motor fault diagnosis based on neuropredictors and wavelet signal processing", *IEEE/ASME Transactions on Mechatronics*, Vol. 7, No. 2, pp. 201-219, 2002.

[20] B. Li, M. Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based motor rolling bearing fault diagnosis", *IEEE Transactions on Industrial Electronics*, Vol. 47, pp. 1060 - 1069, 2000.

[21] B. Liu, S. Riemenschneider, and Y. Xu, "Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum," *Mechanical Systems and Signal Processing*, Vol. 20, pp. 718-734, 2006.

[22] K. A. Loparo, M. L. Adams, W. Lin, F. M. Abdel-Magied, and N. Afshari, "Fault Detection and Diagnosis of Rotating Machinery", *IEEE Transactions on Industrial Electronics*, Vol. 47, No.5, pp. 1005 - 1014, 2000.

[23] J. Ma and D. He, "Development of Integrated Prognostics: An Application to Hybrid Ceramic Bearing Life Prediction", *Proceedings of the 2010 Second IITA International Joint Conference on Artificial Intelligence*, Shanghai, China, December 25 – 26, 2010.

[24] D. Mba, "The use of acoustic emission for estimation of bearing defect size," *Journal of Failure Analysis and Prevention,* Vol. 8, pp. 188-192, 2008.

[25] S. Niizeki, "Ceramic Bearing for Special Environments", *NSK Journal of Motion & Control*, No. 8, pp. 17 - 22, 2000.

[26] H. Ohta and K. Kobayashi, "Vibrations of hybrid ceramic ball bearings", *The Japan Society of Mechanical Engineers*, Vol. 61, No. 589, pp. 4736–4743, 1995.

[27] J. R. Stack, T. G. Habetler, and R. G. Harley, "Bearing fault detection via autoregressive stator current modeling", *IEEE Transactions on Industry Applications*, Vol. 40, pp. 740 - 747, 2004.

[28] M. Rhoads and M. Bashyam, "Engine Hybrid Ceramic Bearings", *WL–TR–95–4081*, Wright Patterson AFB, OH, 1994.

[29] J. A. Rosero, L. Romeral, J. A. Ortega, and E. Rosero, "Short-circuit detection by means of empirical mode decomposition and Wigner–Ville distribution for PMSM running dnder dynamic dondition", *IEEE Transactions on Industrial Electronics*, Vol. 56, No. 11, pp. 4534 – 4547, 2009.

[30] B. Samanta, K. Al-Balushi, and S. Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection", *Engineering Applications of Artificial Intelligence*, Vol. 16, pp. 657 - 665, 2003.

[31] P. Samuel and D. Pines, "A review of vibration-based techniques for helicopter transmission diagnostics," *Journal of Sound and Vibration,* Vol. 282, pp. 475-508, 2005.

[32] B. Sreejith, A. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks", in *The Third IEEE International Conference on Industrial and Information Systems*, Kharagpur, India, pp. 1- 6, 2008.

[33] J. Sun and A. Kabán, "A fast algorithm for robust mixtures in the presence of measurement errors", *IEEE Transactions on Neural Networks*, Vol. 21, No. 8, pp. 1206 – 1220, 2010.

[34] H. Takebayashi, "Bearings for Extreme Environments - Part 3, Basic Performance of Ceramic (Silicon Nitride) Bearings", *KOYO Engineering Journal English Edition*, No. 158E, 2001.

[35] H.A. Talebi, K. Khorasani, and S. Tafazoli, "A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem", *IEEE Transactions on Neural Networks*, Vol. 20, No. 1, pp. 45 – 60, 2009.

[36] C. Tan, P. Irving, and D. Mba, "A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears," *Mechanical Systems and Signal Processing*, Vol. 21, pp. 208-233, 2007.

[37] N. Tandon and B. Nakra, "Defect detection in rolling element bearings by acoustic emission method," *Journal of Acoustic Emission*, Vol. 9, pp. 25-28, 1990.

[38] T. Toutountzakis, C. Tan, and D. Mba, "Application of acoustic emission to seeded gear fault detection," *NDT & E International,* Vol. 38, pp. 27-36, 2005.

[39] A. A. Wereszczak, W. Wang, Y. Wang, M. K. Hadfield, W. T. P. Kirkland, and M. Jadaan, "Rolling Contact Fatigue of Ceramics", *ORNL/TM-2006/521 Report*, 2006.

[40] S. Wu and T. W. S. Chow, "Induction machine fault detection using SOM-based RBF neural networks", *IEEE Transactions on Industrial Electronics*, Vol. 51, No. 1, pp. 183 - 194, 2004.

[41] R. Yan and R. Gao, "Hilbert–Huang transform-based vibration signal analysis for machine health monitoring," *IEEE Transactions on Instrumentation and Measurement*, Vol. 55, pp. 2320 - 2329, 2006.

David He received his B.S. degree in metallurgical engineering from Shanghai University of Technology, China, MBA degree from The University of Northern Iowa, and Ph.D. degree in industrial engineering from The University of Iowa in 1994. Dr. He is Professor and Director of the Intelligent Systems Modeling & Development Laboratory in the Department of Mechanical and Industrial Engineering at The University of Illinois-Chicago. Dr. He's research areas include: machinery health monitoring, diagnosis and prognosis, complex systems failure analysis, quality and reliability engineering, and manufacturing systems design, modeling, scheduling and planning.

Ruoyu Li received his B.S. degree in automatic control and M.S. degree in control theory and control engineering from The Guilin Institute of Electronic Technology, Guilin, China, in 2002 and 2005, respectively. He is currently a candidate toward the Ph.D. degree in industrial engineering in the Department of Mechanical and Industrial Engineering at The University of Illinois-Chicago. His research interests include: nonlinear time-series analysis, rotating machinery condition monitoring, diagnosis and prognosis, sensor networks, and mechatronic systems design.

Junda Zhu received his B.S. degree in mechanical engineering from Northeastern University, Shenyang, China and M.S. degree in mechanical engineering from The University of Illinois-Chicago, 2009. He is currently working toward the Ph.D. degree in industrial engineering in the Department of Mechanical and Industrial Engineering at The University of Illinois-Chicago. His research interests include: lubrication oil condition monitoring and degradation analysis, physics based and data mining based full ceramic bearing diagnostics and prognostics, and AutoCAD/FE and design analysis.

Mikhail Zade received his B.S. degree in mechanical engineering from Nagpur University, India, 2009 and M.S. degree in industrial engineering from The University of Illinois-Chicago, 2011. His research interests include: data mining based full ceramic bearing diagnostics.