IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ADVANCES IN PROGNOSTICS AND SYSTEM HEALTH MANAGEMENT

Received February 20, 2019, accepted March 3, 2019, date of publication March 6, 2019, date of current version April 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2903355

A Wind Turbine Bearing Performance Evaluation Method Based on Similarity Analysis of Fuzzy k-Principal Curves in Manifold Space

HONGJI REN[1,2], AIJUN YIN[1,2], (Member, IEEE), QUAN ZHOU[3], JIANG LI[1,2], AND YIHUA HU[4], (Senior Member, IEEE)

[1]State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, China
[2]College of Mechanical Engineering, Chongqing University, Chongqing 400044, China
[3]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996 USA
[4]Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3BX, U.K.

Corresponding author: Aijun Yin (aijun.yin@cqu.edu.cn)

This work was supported in part by the Key Projects of Chongqing's major theme of artificial intelligence technology innovation under Grant cstc2017rgzn-zdyfx007, and in part by the National Natural Science Foundation of China under Grant 51374264.

ABSTRACT Condition monitoring (CM) is widely used in wind turbines (WTs) to reduce operation and maintenance (O&M) costs. Bearings are crucial components in WT and many bearing CM approaches have focused on vibration analysis. Statistical theory and artificial intelligence-based WT bearings evaluation methods require mass data for training, which makes the detection of incipient failures barely possible. In this paper, a WT bearing performance evaluation method is proposed based on the similarity analysis of fuzzy *k*-principal curves (FKPCs) in manifold space. For a start, 38 features are extracted from bearing vibration signals to constitute high-dimensional feature matrices. The feature matrices for the healthy samples and the samples to be evaluated are then transformed into 3-D space. Afterward, the FKPCs are extracted and the similarities among the curves of samples are calculated based on the Hausdorff distance to evaluate the performance of the bearing. Bearing degradation experiments are investigated to verify the efficiency of the proposed method. The results indicate that the proposed FKPC method can portray the degradation trend of bearings accurately with the capability of detecting incipient failures. The proposed method can be applied in the case of small-size training samples with stable performance.

INDEX TERMS Wind farms, condition monitoring, fault diagnosis, prognostics and health management.

I. INTRODUCTION

With the exhaustion of traditional energy resources, more people are turning their attention to wind energy since it is clean and renewable [1]. Energy market is sensitive to costs and wind energy is struggling to expand its share. Wind turbines (WTs) convert the kinetic energy of wind to electrical power. The operation and maintenance (O&M) costs of WTs take up a large proportion in the overall energy generation cost [2]. It will make wind energy more competitive in the energy market if the O&M costs are reduced. A larger number of efforts have been devoted to optimizing the mechanical design of WTs for higher efficiency and durability [3]. It is also important to evaluate the performance of WTs accurately so that the faulty parts can be repaired prior to the occurrence of more serious issues [4]. There are many approaches that aim at reducing WTs' O&M costs such as condition monitoring (CM), non-destructive testing (NDT) [5], fault diagnosis (FD) [6], etc.

A typical WT mainly consists of the blades, bearings, gearbox, generator, shaft, yaw system and tower. Blades rotate around the horizontal axis and the kinetic energy is transferred to the generator via the shaft, bearings and gearbox. There are also bearings in the generator and gearbox. The amount of electrical energy generated by a WT depends on the size of the turbine and the wind speed. Larger blades and stronger wind will lead to more power generation as well as more abrasion and damage to the components. WTs are inevitably subject to failure after long-time service [7]. WT is a typical mechatronics system, in which functional components are structurally and electrically connected. The vulnerable components in WT are easy to break down under severe operational environments, in which case the O&M costs are increased. Failure causes and modes for components in a

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

TABLE 1. Failure causes and modes for components in a WT system.

Component	Failure causes	Failure modes
Generator	Over speed, dirt of corrosion, excessive vibration, overload, excessive temperature, ambient conditions, debris	Core insulation, slip ring failure, brushgear failure, electrical trip, commutator failure, stator & rotor mechanical integrity failure
Blade	Lightning strike, fatigue cycle, wind gust, moisture absorption	Delamination, fatigue crack, splitting, fracture, disbanding, impact
Gearbox	Misalignment, fatigue cycle, starting torque, lack of oil, unexpected loads, resonant vibration, surface wear	Bearing seize, teeth stripped
Bearing	Over speed, debris, excessive vibration, overload, excessive temperature, ambient conditions	Fatigue, plastic deformation, abrasion, electro erosion, crack, corrosion

WT system are illustrated in Table 1 according to previous survey [8].

From Table 1 we can see that the problems of WTs are mainly caused by failures of the blades, generator, bearings and gearbox. Extensive efforts have been devoted to evaluating the performance of WT components [9]–[12]. Bearings are important parts in WTs. According to [13], the majority of wind turbine gearbox failures (about 70%) are caused by faults in the bearings. Besides, the performance of the generator is also related to the bearings. Vibration analysis is an efficient method in the performance evaluation of rolling components and has been widely applied in WT bearings monitoring [14]. Generally, there are two popular branches for WT bearings vibration signal analysis: the conventional time and frequency analysis methods and the novel data mining approaches [15]. A number of time-frequency based bearing evaluation methods such as sparse decomposition, Hilbert-Huang Transform (HHT), empirical mode decomposition (EMD), local mean decomposition (LMD) and (VMD) have been proposed. However, the above methods have their respective limitations, and the estimation accuracy generally depends on the parameters settings [16], [17]. Therefore, novel data mining approaches based on statistical theory and artificial intelligence have been brought forward, such as hidden Markov models (HMM) [18], deep belief networks (DBN) [19], artificial neural network (ANN) [20], and fuzzy formalisms [21]. Although great developments have been made in both the theoretical research and practical applications, statistical and artificial intelligence based methods require masses of data for model training. Moreover, since the evaluation results are very sensitive to the parameters of the models while the optimal parameter sets are not always available, the above mentioned methods are generally not capable of detecting incipient failures, especially when the data is limited.

The motivation of this study is to present a WT bearings performance evaluation method with the capability of detecting incipient failures. The feature vectors extracted from WT bearings vibration signals are usually high-dimensional, nonlinear, information-redundant and mutually coupled [22]. Manifold learning is a dimension reduction method that can extract the low-dimensional manifold structure of target data in the original observation space [23]. As one of the classical manifold learning algorithms, Laplacian eigenmaps has been successfully applied on dimension reduction and data representation [24]. After dimension reduction, an imaginary manifold curve can be extracted to reflect the condition of the bearings by applying geometric interpolation [25]. The Hausdorff distance, which is generally used to evaluate the similarity among curves [26], can be used to calculate the distances among manifold curves under different conditions thus to reflect the degradation process. To sum up, the detection of bearing incipient failures can be realized through an optimized combination of dimension reduction method, manifold curve method and similarity comparison algorithm, which has not been accomplished yet.

In this study, a WT bearing performance evaluation method based on the similarity analysis of fuzzy K-principal curves (FKPC) in manifold space is presented. The proposed method is sensitive to incipient failures, and in the meanwhile large volumes of data is not necessary in the training process. The rest of this paper is organized as follows: Feature extraction and Laplacian eigenmaps are introduced in Section II. In Section III, a novel FKPC method is proposed and the Hausdorff distance is utilized to assess the similarities of curves. Experiments and results are demonstrated in Section IV. Finally, the conclusion is presented in Section V.

II. FEATURE EXTRACTION AND LAPLACIAN EIGENMAPS

A. FEATURE EXTRACTION

Vibration signals are widely used in the evaluation of bearings. In this paper, features extracted from multiple domains are employed to create high dimensional feature matrices for dimension reduction. Time-domain features, which are intuitive and intelligible, constitute the raw data of the bearing running state. Frequency domain features are used to describe the variations in the frequency band from the view of signal spectrum and spectrum energy distribution. 24 time and frequency features are utilized in this paper according to [27], shown in Table 2 and Table 3.

A wavelet packet transformation is deduced based on decomposition of the low frequency portion by a discrete wavelet transformation. Then a more thorough decomposition of the high-frequency portion of the signal without redundancy or oversight can be completed, which provides better local time-frequency analytic ability. Wavelet packet node energy is one of the most widely used features in mechanical fault analysis. The energy value at each node can be derived by:

$$E_{jk} = \sum_{m=1}^{N} x_{jm} \tag{1}$$

TABLE 2. Features in the time-domain.

Symbol	Equation	Symbol	Equation
p_1	$\frac{\sum_{n=1}^{N} x(n)}{N}$	p_7	$\frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-p_1)}{p_2^{\ 3}}$
p_2	$\sqrt{\frac{\sum_{n=1}^{N}(x(n)-p_1)^2}{N-1}}$	p_8	$\frac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-p_1)}{p_2^{\ 4}}$
p_3	$\left(\frac{\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}}{N}\right)^2$	p_9	$\frac{p_5}{p_4}$
p_4	$\sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}$	p_{10}	$\frac{p_5}{p_3}$
p_5	$\max\lvert x(n)\rvert$	p_{11}	$\frac{p_5}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
p_6	$\frac{p_4}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$		

$x(n)$ is the sequence of time domain signal, n=1, 2, ..., N, N is the total number of samples.

TABLE 3. Features in the frequency-domain.

Symbol	Equation	Symbol	Equation
p_{12}	$\frac{\sum_{a=1}^{A} s(a)}{A}$	p_{19}	$\sqrt{\frac{\sum_{a=1}^{A} f_a^{\ 4} s(a)}{\sum_{a=1}^{A} f_a^{\ 2} s(a)}}$
p_{13}	$\sqrt{\frac{\sum_{a=1}^{A}(s(a)-p_{12})}{A-1}}$	p_{20}	$\frac{\sum_{a=1}^{A} f_a^{\ 2} s(a)}{\sqrt{\sum_{a=1}^{A} s(a)\sum_{a=1}^{A} f_a^{\ 4} s(a)}}$
p_{14}	$\frac{\sum_{a=1}^{A}(s(a)-p_{12})}{A\left(\sqrt{p_{13}}\right)^3}$	p_{21}	$\frac{p_{17}}{p_{16}}$
p_{15}	$\frac{\sum_{a=1}^{A}(s(a)-p_{12})^4}{Ap_{13}^{\ 2}}$	p_{22}	$\frac{\sum_{a=1}^{A}(f_a-p_{16})^3 s(a)}{Ap_{17}^{\ 3}}$
p_{16}	$\frac{\sum_{a=1}^{A} f_a s(a)}{\sum_{a=1}^{A} s(a)}$	p_{23}	$\frac{\sum_{a=1}^{A}(f_a-p_{16})^4 s(a)}{Ap_{17}^{\ 4}}$
p_{17}	$\sqrt{\frac{\sum_{a=1}^{A}(f_a-p_{16})^2 s(a)}{A}}$	p_{24}	$\frac{\sum_{a=1}^{A}(f_a-p_{16})^{1/2} s(a)}{Ap_{17}^{\ 1/2}}$
p_{18}	$\sqrt{\frac{\sum_{a=1}^{A} f_a^{\ 2} s(a)}{\sum_{a=1}^{A} s(a)}}$		

$s(a)$ is signal spectrum, a is spectral line. f_a is the frequency of a-th spectral line, A is the total number of spectral lines. a=1, 2, ..., A.

where E_{jk}, m and x_{jm} are the corresponding signal energy for the reconstructed signal of j-th band in k-th layer, the discrete node and the amplitude of the discrete node, respectively.

The wavelet packet node energy eigenvector is defined as:

$$e = \left\{E_{j0}, E_{j1}, \ldots, E_{jl}\right\}/E \tag{2}$$

where $l = 2^j - 1$, $E = \sum_{k=1}^{l} E_{jk}$

B. LAPLACIAN EIGENMAPS FOR MANIFOLD SPACE PROJECTION

The original high dimensional feature matrices are usually nonlinear, information-redundant and mutually coupled. It is necessary to generate an essential representation of the original feature matrices by applying space conversion. The high dimensional feature matrices (the D dimension) are actually in a low dimensional manifold (the L dimension) and the manifold structure retains the geometrical characteristics of the feature matrices. Laplacian eigenmaps is a classical manifold learning method that can obtain a low dimensional representation of the datasets and maintain local features with robust performance when the samples have outliers [28].

Given a dataset $\boldsymbol{X} = [x_1 \ldots, x_{\mathrm{M}}]$ of M points in $\mathbf{R}^{\mathrm{D}}$, the purpose of Laplacian eigenmaps is to find an optimal embedding Ω for the manifold. Suppose that $Y = [y_1 \ldots, y_{\mathrm{M}}] \epsilon \mathbf{R}^{\mathrm{L}}$ is the low-dimensional representation of $\boldsymbol{X}$. $\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X})$, where $\boldsymbol{F}$ is the embedding map.

$\boldsymbol{F}$ is used to ensure that the distance relations among data points are maintained during dimension reduction. $\boldsymbol{F}$ can be obtained through minimizing the following objective function:

$$\begin{cases} \operatorname{argmin} \int \left\| \nabla \boldsymbol{F}(X)^2 \right\| \\ \|\boldsymbol{F}\|_{\Pi^2(M)} = 1 \end{cases} \tag{3}$$

The minimization of (3) can be transformed to a generalized eigenvalue problem:

$$\Gamma \boldsymbol{F} = -div \nabla(\boldsymbol{F}) \tag{4}$$

where Γ is the Laplace Beltrami operator. Let $f_0, f_1, \ldots, f_{\mathrm{M}}$ be the solutions of equation (4) ordered according to their eigenvalues $0 = \lambda_0 \leq \lambda_1 \ldots \leq \lambda_{\mathrm{M}}$, and the optimal embedding map can be described as $\boldsymbol{F} = [f_1, f_2, \ldots, f_{\mathrm{M}}]^{\mathrm{T}}$. Then a high dimensional space can be projected into low dimensional manifold space by $\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X}) = [f_1(\boldsymbol{X}), f_2(\boldsymbol{X}), \ldots, f_{\mathrm{M}}(\boldsymbol{X})]$.

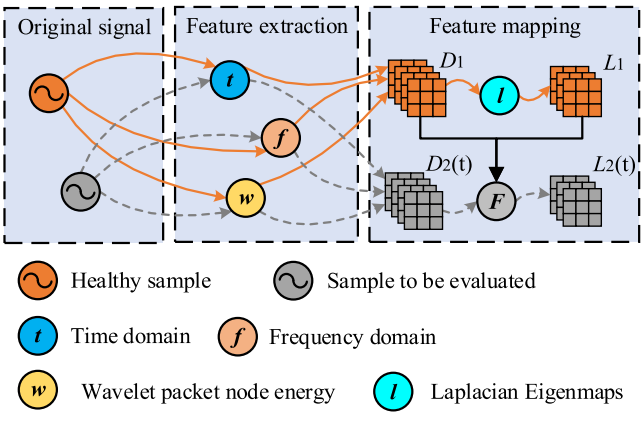

FIGURE 1. Flow chart of space conversion of rolling bearing vibration signals.

The flow chart of space conversion of bearing vibration signals is illustrated in Fig.1. Time domain features, frequency domain features and wavelet packet node energy

features constitute the high-dimensional feature matrices. D_1 and D_2(t) denote the high-dimensional feature matrices for the healthy samples and the samples to be evaluated, respectively. Firstly, Laplacian eigenmaps is employed to transform D_1 to the low dimensional manifold space matrix of the healthy samples L_1, and in the meanwhile the project map $\boldsymbol{F}$ is derived. $\boldsymbol{F}$ is then utilized to transform D_2(t) to the low dimensional manifold space matrix of the samples to be evaluated L_2(t). Further comparison between L_1 and L_2(t) will be conducted for the bearing performance evaluation.

III. SIMILARITY ASSESSMENT OF FUZZY k-PRINCIPAL CURVES

A. EXTRACTION OF FUZZY k-PRINCIPAL CURVES

The feature matrix after Laplacian eigenmaps is the low-dimensional expression of the original high-dimensional target data. The principal curve aims to represent exactly the distribution of target data in low-dimensional space and it can be regarded as a nonlinear generalization of PCA [29]. To calculate the master curve from the dataset, a curve satisfying certain optimization objectives and constraints is selected from a set of curves. The principal curve is essentially a one dimensional manifold embedded in Euclidean space [30].

Given the probability density of a random variable $\boldsymbol{B} = (B_1, B_2, \ldots, B_n)$, if a curve $c\,(s)$ passing through $\boldsymbol{B}$ satisfies the following expression:

$$c\,(s) = E\,(B\,|s_b\,(b) = s\,) \tag{5}$$

where $c(s)$ is one of the principal curves of $\boldsymbol{B}$, $s_b(b) = \sup\{s||b - c(s)|| = \inf||bc(\tau)||\}$ is the projection value from data points B to the points in curve $c(s)$ and E is the mean

The k-principal curve, which is also known as the principal curve with length constraint, is the only existing principal curve. Based on [31], a new method of extracting the principal curve by using fuzzy k-means clustering is proposed in this paper. Fuzzy k-means is a classical dynamic clustering algorithm that constantly minimizes the square error and updates the membership degree of each data sample point to the cluster center. Therefore, the cluster with the highest membership degree indicates the classification of the data samples.

For n samples with q variables, after normal transformation and standardization, a fuzzy matrix of the initial m groups of each sample is given as:

$$U = \begin{bmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{n1} & \cdots & u_{nm} \end{bmatrix} i = 1, 2 \cdots n; j = 1, 2 \cdots m \tag{6}$$

where u_{ij} is the membership degree of sample i for class j, $u_{\mathrm{ij}} \epsilon [0, 1]$, $\sum u_{\mathrm{ij}} = 1$. Let v_{j} be the initial cluster center of class j

$$V = [v_1, v_2 \cdots v_m] \tag{7}$$

where $v_{\mathrm{j}} = [v_{\mathrm{j1}}, v_{\mathrm{j2}} \ldots, v_{\mathrm{jq}}]$. Define an objective function $J_z(UV)$:

$$J_z\,(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(u_{ij}\right)^z \left\|x_i - v_j\right\|^2 \tag{8}$$

where x denotes a point with q variables, z is the smoothing parameter, $1 \leq z \leq 5$ and z is usually set to 1 This problem can be attributed to the extreme value of the objective function on the condition of $\sum u_{\mathrm{ij}} = 1$. When $m > 1$, $x_{\mathrm{i}} \neq v_{\mathrm{j}}$, which can be proved as follows by using Lagrange multiplication:

$$\begin{aligned} u_{ij} &= 1/\sum_{j=1}^{m} \left(\left\|x_i - v_j\right\| / \left\|x_i - v_1\right\|\right)^{2/(z-1)} \\ v_j &= \sum_{i=1}^{n} u_{ij}^z x_i / \sum_{i=1}^{n} u_{ij}^z \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m \end{aligned} \tag{9}$$

The function of k-principal curve is minimizing (10) continuously. Then, a 2-opt algorithm is used to optimize the Hamilton path formed by line segments until they are connected to a smooth curve [26].

$$min \sum_{i=1}^{k} \sum_{x \in V_i} d\,(x, s_i)^2 \tag{10}$$

where s_{i} denotes the i-th line segment, $d(x, s_{\mathrm{i}})$ is the distance from x to s_i, and k is the total number of line segments. $V_{\mathrm{i}} = \left\{ x \in X_n | i = \begin{matrix} \arg\min d\,(x, s_i) \\ j \end{matrix} \right\}$. X_{n} is a set of n samples from R. V_i is the Voronoi Region.

Therefore, the method of extracting the principal curve by using fuzzy k-means clustering can be summarized as follows:

Step1: Use Fuzzy k-means clustering algorithm to divide the original data points into m classes;

Step2: Iterate continuously to minimize (10) and obtain k line segments;

Step3: Optimize the Hamilton path to connect k line segments to a smooth curve.

B. CURVE SIMILARITY

Judging curve similarity is one of the foremost problems of computer graphics and pattern recognition. Currently, the main judgment methods are the eigenvalue method and the similarity functions definition-based method. The similarity functions definition-based methods are superior to the eigenvalue method to a certain extent. The Hausdorff distance is widely used in similarity measurement between two-point sets [32].

The Hausdorff distance is a nonlinear distance, which only calculates the degree of similarity (maximum distance) between two points sets without establishing the corresponding relationship between the points.

For points sets $A = \{a_1, a_{2,} \ldots, a_{\mathrm{p}}\}$ and $B = \{b_1, b_2, \ldots, b_{\mathrm{q}}\}$, the Hausdorff distance between A and B is:

$$H_{\mathrm{MHD}}\,(A, B) = max\,\{h_{\mathrm{MHD}}\,(A, B)\,,\, h_{\mathrm{MHD}}\,(B, A)\} \tag{11}$$

where

$$h_{\mathrm{MHD}}(A,B) = \max_{a_i \in A} \min_{b_j \in B} \left\| a_i - b_j \right\| \tag{12}$$

$$h_{\mathrm{MHD}}(B,A) = \max_{b_j \in B} \min_{a_i \in A} \left\| b_j - a_i \right\| \tag{13}$$

Note that $||a_i - b_j||$ is the Euclidean distance between the points a_i and b_j

C. THE ASSESSMENT OF BEARING PERFORMANCE BY USING THE FKPC

In this paper, firstly, the original high dimensional feature matrices are transformed to those in three dimensional space by using Laplacian eigenmaps. Secondly, the fuzzy k-principal curves for the healthy samples and the samples to be evaluated are extracted, respectively. Finally, the similarities are calculated by applying the Hausdorff distance. The algorithm process is illustrated in Table 4.

TABLE 4. The algorithm process of the FKPC.

Assessing bearing performance by using the FKPC
Input: bearing vibration signals. *Step*1: Extracting the feature matrices, The healthy samples: D_1, the samples to be evaluated: D_2(t). *Step*2: Space conversion based on Laplacian eigenmaps. Use (4) to obtain low-dimensional representation of D_1, and project map $\boldsymbol{F}$, then use $\boldsymbol{F}$ to assess the low-dimensional representation of D_2(t). *Step*3: fuzzy k-principal curve extraction. Calculate the fuzzy k-principal curves of healthy samples and samples to be evaluated, which are f_1 and f_2(t), respectively. *Step*4: Similarity assessment, $Dis(t) = H_{\mathrm{MHD}}(f_1, f_2(t))$.

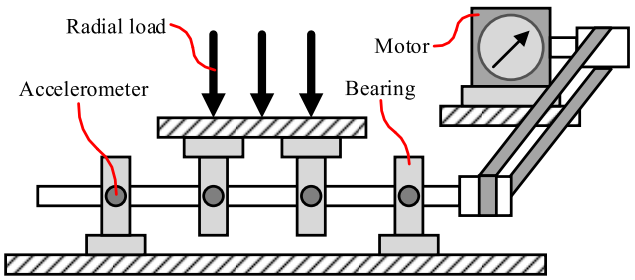

FIGURE 2. The experiment setup for the accelerated degradation of the rolling bearing.

IV. EXPERIMENTS AND RESULTS

A. LABORATORIAL BEARING DEGRADATION EXPERIMENT SETUP

The experiment setup for the accelerated degradation of the rolling bearing is shown in Fig.2. Four ZA-2115 double row roller bearings installed on the test bench are driven by the AC motor through a friction pulley with a rotation speed of 2000RPM. A 26700N radial load is applied in the test bench beam to accelerate the degradation, and the vibration data is collected by an NI DAQ Card 6062E with a sampling frequency of 20kHz [34]. The experiment is stopped when the amount of accumulated debris exceeds a certain threshold. There are four channels, with channels 1 to 4 corresponding to bearings 1 to 4, respectively. The dataset of bearing 1 is selected for further analysis.

B. LABORATORIAL BEARING DEGRADATION EXPERIMENT RESULTS

24 time and frequency features are extracted from the raw vibration signal. We then obtain 14 sub-band node energy values through 3-layer decomposition by using a db3 wavelet function of the wavelet packet transform. Thus, 38 features constitute the original feature sets. Fig.3 shows the variation trends of 4 different sensitive features. We can see from Fig.3 that the performance varies from feature to feature. Therefore, the feature extraction process can be arduous and frustrating if the process is performed manually. Fig.3 also shows that the bearing is in healthy state before 4000min. In this paper, healthy samples are selected from data in 0-1000min and the entire data sets are the samples to be evaluated.

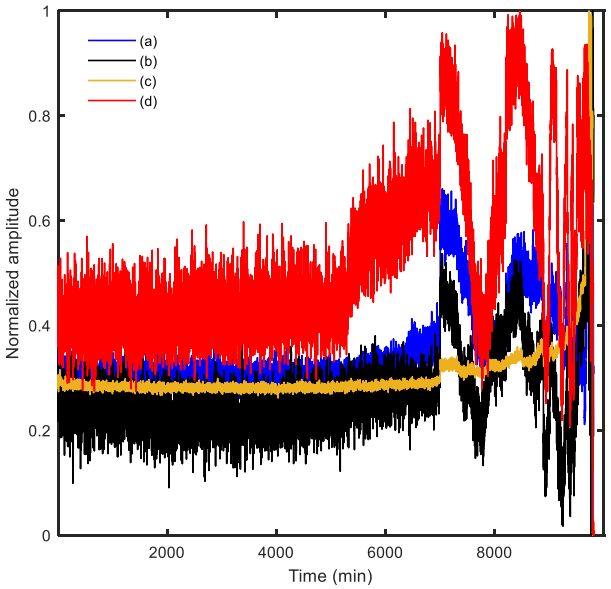

FIGURE 3. Variation trend of 4 features: (a) waveform index of time domain; (b) first order gravity center of frequency domain; (c) second order gravity center of frequency domain; (d) energy of node (3,7) in wavelet packet domain.

We combine the time domain features, frequency domain features and wavelet packet energy features to constitute the high dimensional feature matrices to comprehensively reflect the running state of the bearing. The feature matrix for the healthy bearing state is transformed to three-dimensional space with preserved projection map information by using Laplacian eigenmaps. The feature matrices of the samples to be evaluated are then transformed to three-dimensional space according to the projection map.

The fuzzy k-principal curves are extracted from the feature sets after space conversion The similarities among the

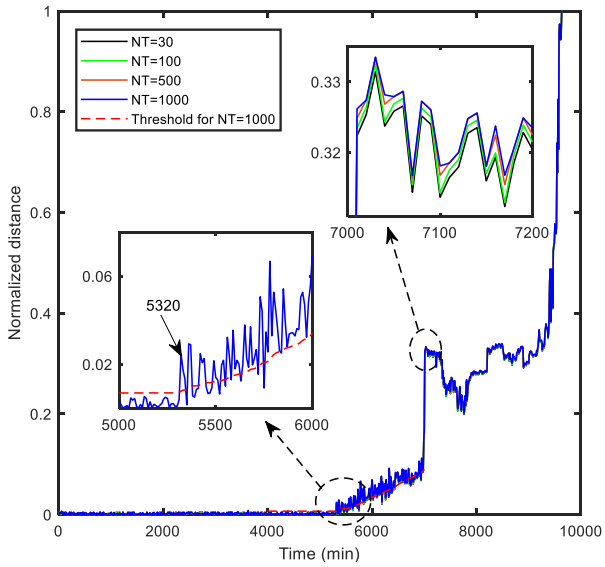

FIGURE 4. The bearing degradation curve by using the FKPC "NT" denotes the size of training samples.

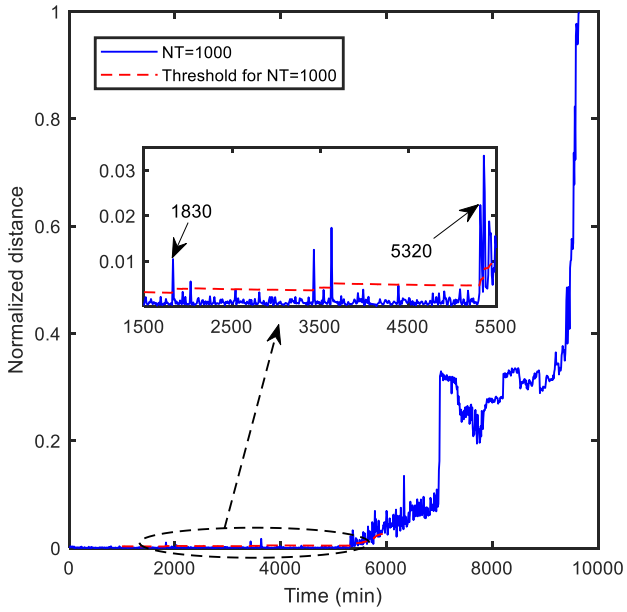

FIGURE 5. The bearing degradation curve by directly calculating the Hausdorff distance after Laplacian eigenmaps. Since the influence of NT is much smaller than it in Fig.4, here we focus on the result under 1000 training samples.

principal curves are calculated by using the Hausdorff distance based method. Fig.4 illustrates the bearing degradation curve by using the FKPC. "NT" denotes the size of training samples. The minimum sample size is set to 30 to take into consideration the poor performance of the HMM under small sample sizes. In Fig.4, the degradation curve is stable before 5000min, indicating that the bearing is running under a normal condition. The incipient failure starts at 5320min, which means that the deviation from the normal operating condition is gradually increasing. The fluctuation at 7010min shows that the performance of the bearing is worsen. And the total failure comes at about 9670min. The FKPC generally reproduces the degradation trend of the bearing and the performance of the FKPC is stable despite the variation of training sample numbers. The bearing degradation curve by directly calculating the Hausdorff distance after Laplacian eigenmaps is illustrated in Fig.5. We can see that Fig.5 can also portray the degradation process of the bearing. However, further comparison between Fig.5 and the proposed method in Fig.4 shows that the curve in Fig.5 has obvious fluctuation. The degradation curve by applying FKPC is stable before 5000min and the incipient failure starts at 5320min. While in Fig.5 the incipient failure starts at 1830min, which is inaccurate according to the dataset itself. Therefore, directly analyzing the Hausdorff distance on the feature matrix after Laplacian eigenmaps is vulnerable to noise or abnormal data. The FKPC, on the other hand, can reduce above influence and thus to more accurately represent the degradation process of the bearing.

C. COMPARISON WITH OTHER DIMENSION REDUCTION METHODS

In this paper, high dimensional feature matrices are expressed in three-dimensional space to reflect the running state of bearings. The efficacy of the three-dimensional space is therefore affected by the dimension reduction methods. In this section, the performance of PCA and Isomap based dimension reduction methods are also investigated.

The k-principal curves are extracted from the feature sets after dimension reduction, and the similarities among the principal curves of the healthy samples and the samples to be evaluated are then calculated by using the Hausdorff distance based method. We use the 3σ threshold to detect incipient failure according to [33]. The bearing begins to fail if the degeneration curve starts exceeding the threshold evidently. Fig.6 illustrates the bearing degradation curves by using the Laplacian eigenmaps, PCA and Isomap. Both the Laplacian eigenmaps and PCA can portray the degradation trend of the bearing while the Isomap results in distance fluctuations with unobvious tendencies. This can be explained by the fact that the Isomap method is not capable of dealing with manifolds with larger internal curvatures. Further comparison shows that the incipient failure detection results of the Laplacian eigenmaps and PCA are 5320min and 5490min, respectively. Therefore, the performance of the Laplacian eigenmaps is the best among the three methods.

D. COMPARISON WITH OTHER DISTANCE SIMILARITY ASSESSMENT APPROACHES

The Hausdorff distance is widely used in the evaluation of curve similarity. The approaches by applying the Euclidean distance and Mahalanobis distance are also popular. Fig.7 illustrates the bearing degradation curves by using the Hausdorff distance, Euclidean distance and Mahalanobis distance. We can see that the performance of Euclidean

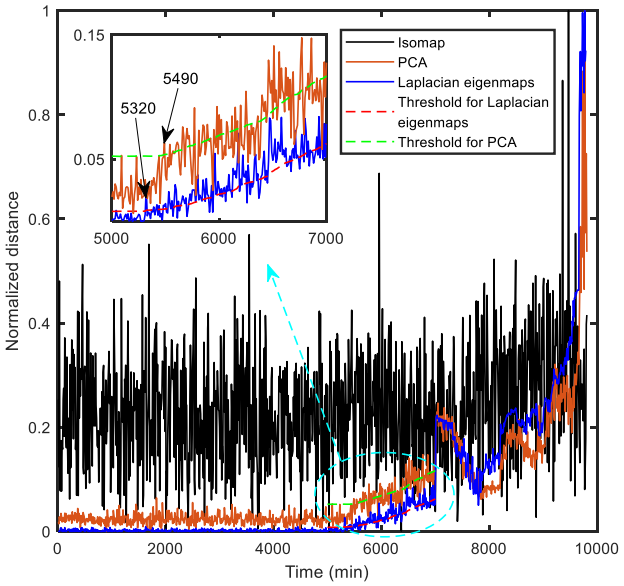

FIGURE 6. The bearing degradation curves by using Laplacian eigenmaps, PCA and Isomap.

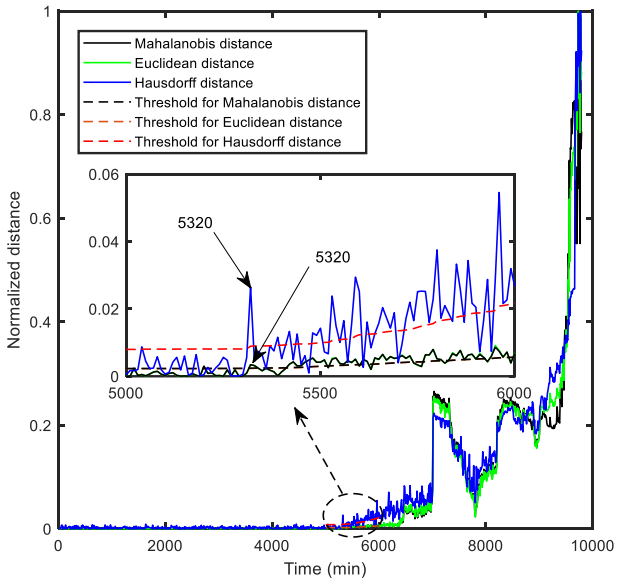

FIGURE 7. The bearing degradation curves by using the Hausdorff distance, Euclidean distance and Mahalanobis distance.

distance and Mahalanobis distance are similar, and all of the three methods are able to reproduce the degradation process of the bearing. As for the ability to detect incipient failure, the Hausdorff distance exceeds the others since the failure symptom at 5320min as well as other time points is more significant.

E. COMPARISON WITH THE HMM AND THE DBN

Based on the high-dimensional feature matrices, the bearing degradation curves by using the Gaussian mixture HMM and DBN are illustrated in Fig.8 and Fig.9, respectively.

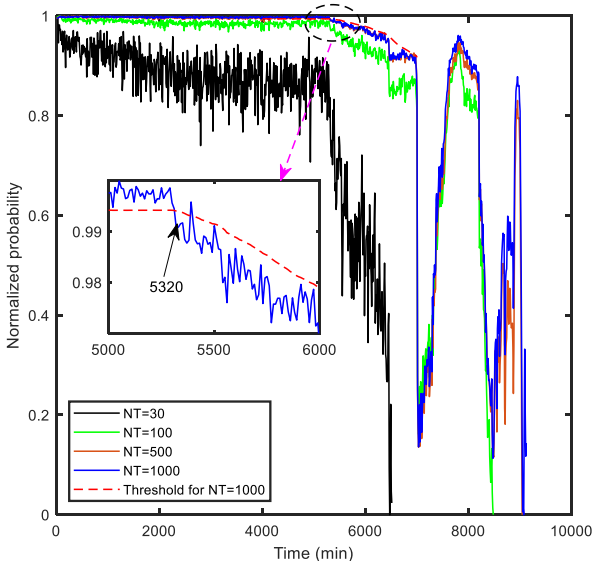

FIGURE 8. The bearing degradation curve by using the HMM.

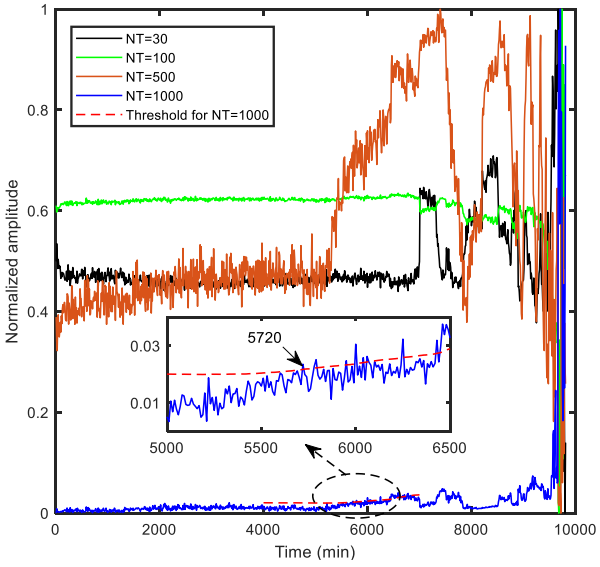

FIGURE 9. The bearing degradation curve by using the DBN.

Fig.8 indicates that the performance of the HMM becomes better as the number of samples increases. The key parameters of a HMM include initial probability distribution, transition probability distribution and observation probability distribution. For a WT bearing monitoring case where the observations are continuous, the parameters of the HMM is more complicated due to the Gaussian distribution. The same pattern can also be found in the performance of the DBN, shown in Fig.9. The detailed comparison among the 3 methods is summarized in Table 5. It is clear that for statistics based bearing performance evaluation methods such as the HMM and DBN, the optimized parameters are not always available and the methods may fail to reflect the

TABLE 5. The comparison among the FKPC, the HMM, and the DBN.

Comparison indicator	**FKPC**	HMM	DBN
Time of incipient failure (min)	**5320**	**5320**	5720
Performance under different training sample sizes	**Stable**	Unstable	Unstable

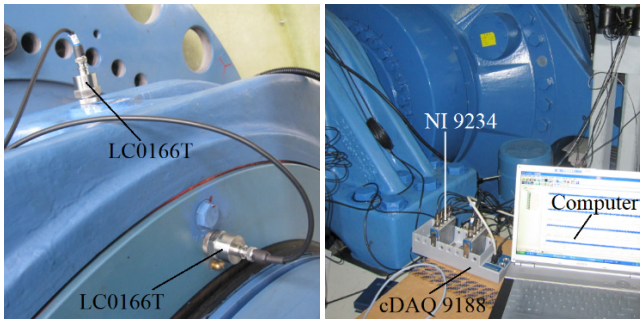

FIGURE 10. The experiment setup for the actual WT bearing experiment. (a) The acceleration sensors settings for main shaft bearings, (b) The acquisition system set-up.

real conditions. The proposed method is sensitive to incipient failures. Moreover, the proposed method can be applied in the case of small-size training samples and the performance of the proposed method is stable despite the variation of training sample numbers.

F. ACTUAL WT BEARING EXPERIMENT AND DISCUSSION

The purpose of the laboratorial bearing degradation experiment is to rapidly perform the degradation process of the bearing and the operating condition is different with actual situation. Therefore, an actual WT bearing experiment has also been conducted on a wind farm located at Inner Mongolia, China. The actual WT bearing experiment aims to assess the natural degradation trend of the bearing. A 2MW WT manufactured by Baoding Tianwei is employed in the experiment. The vibration signals of the main shaft bearings are collected using acceleration sensors LC0166T. The sensors are connected to a computer through NI 9234 and cDAQ-9188. The wind turbine bearing experimental set-up is illustrated in Fig.10. The wind turbine runs under normal condition and a total of 4 datasets are collected at October of 2011 (T1), May of 2012 (T2), October of 2012 (T3) and April of 2013 (T4). During each acquisition, the main shaft of the WT rotates with a speed of 15.2RPM and the vibration data is collected with a sample frequency of 2.56kHz. The datasets are collected manually on the wind farm with an interval of months, so that the condition variation of bearings can be traced. The occasion of collection also depends on the arrangement of the production schedule. The 4 datasets reflect 4 different conditions of the WT bearing.

Since the number of datasets of the actual WT bearing experiments is much smaller than that in the laboratorial

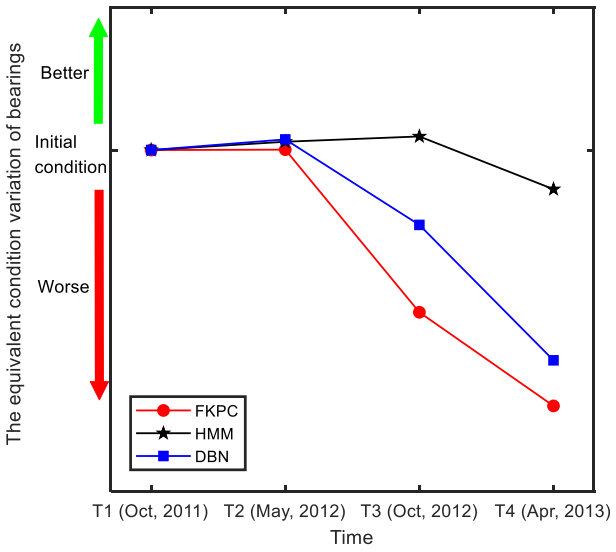

FIGURE 11. The equivalent condition variations of bearings by using the FKPC, the HMM and the DBN.

bearing degradation experiments, we focus on the equivalent condition variation of the bearings instead. The data collected at October of 2011 is selected as the training sample and the condition reference. The 4 datasets are the test samples. The equivalent condition variations of bearings by using the FKPC, HMM and DBN are illustrated in Fig.11. In this case, the condition of the bearings worsens over time. Therefore, the trend by applying the FKPC is generally accurate. The trend by applying HMM shows better condition in T2 and T3, and the trend by applying DBN shows better condition in T2. Both HMM and DBN fail to portray the condition variation of bearings. In conclusion, despite limited data, the FKPC is still able to represent the real condition.

V. CONCLUSION

This paper presents a manifold learning based WT bearing performance evaluation method. The similarity analysis of fuzzy k-principal curves for different bearing conditions is demonstrated by applying the Hausdorff distance. The implementation of the proposed method includes four stages: feature extraction, Laplacian eigenmaps based dimension reduction, k-principal curve estimation, and Hausdorff distance calculation. Experiment analyses have also been conducted and the respective impacts of dimension reduction and distance calculation approaches are discussed. The results indicate that: 1) The proposed FKPC method is able to reflect the overall degradation trend of bearing as well as identify incipient failures; 2) The proposed method can deal with the case of small-size training samples; 3) The performance of the proposed method is stable despite the variation of training sample numbers. Further studies will focus on optimizing the process of fuzzy k-principal curve extraction. The initial number of iterations during the extraction is manually set

in current approaches, which makes the extraction process a little complicated. Therefore, adaptive extraction will be an interesting direction of research to simplify the process.

REFERENCES

[1] X. Zhao, Z. Yan, and X.-P. Zhang, "A wind-wave farm system with self-energy storage and smoothed power output," *IEEE Access*, vol. 4, pp. 8634–8642, Nov. 2016.

[2] Y. Chen, Z. Y. Dong, K. Meng, F. Luo, Z. Xu, and K. P. Wong, "Collector system layout optimization framework for large-scale offshore wind farms," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1398–1407, Oct. 2016.

[3] H. Geng, D. Xu, B. Wu, and G. Yang, "Active damping for PMSG-based WECS with DC-link current estimation," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1110–1119, Apr. 2011.

[4] L. Wang, T.-H. Yeh, W.-J. Lee, and Z. Chen, "Benefit evaluation of wind turbine generators in wind farms using capacity-factor analysis and economic-cost methods," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 692–704, May 2009.

[5] B. Gao, Y. He, W. L. Woo, G. Y. Tian, J. Liu, and Y. Hu, "Multidimensional tensor-based inductive thermography with multiple physical fields for offshore wind turbine gear inspection," *IEEE Trans. Ind. Electron.*, vol. 63, no. 10, pp. 6305–6315, Oct. 2016.

[6] J.-H. Zhong, J. Zhang, J. Liang, and H. Wang, "Multi-fault rapid diagnosis for wind turbine gearbox using sparse Bayesian extreme learning machine," *IEEE Access*, vol. 7, pp. 773–781, Dec. 2018.

[7] Q. Chen, Y. Li, and J. E. Seem, "Bumpless transfer-based inter-region controller switching of wind turbines for reducing power and load fluctuation," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 23–31, Jan. 2016.

[8] C.-P. Malliou, A. D. Karlis, M. G. Danikas, and B. Lloyd, "A short review on the offshore wind turbine generator windings' insulation and the effect of water droplets and salinity," *IEEE Trans. Ind. Appl.*, vol. 52, no. 6, pp. 4610–4618, Nov./Dec. 2016.

[9] Z. Hameed, Y. S. Hong, Y. M. Cho, S. H. Ahn, and C. K. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review," *Renew. Sustain. Energy Rev.*, vol. 13, no. 1, pp. 1–39, Jan. 2009.

[10] S. Zhang, T. L. Jensen, O. Franek, P. C. F. Eggers, C. Byskov, and G. F. Pedersen, "Investigation of a UWB wind turbine blade deflection sensing system with a tip antenna inside a blade," *IEEE Sensors J.*, vol. 16, no. 22, pp. 7892–7902, Nov. 2016.

[11] P. Guo, D. Infield, and X. Yang, "Wind turbine generator condition-monitoring using temperature trend analysis," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 124–133, Jan. 2012.

[12] Y. Qin, J. Zou, and F. Cao, "Adaptively detecting the transient feature of faulty wind turbine planetary gearboxes by the improved kurtosis and iterative thresholding algorithm," *IEEE Access*, vol. 6, pp. 14602–14612, 2018.

[13] S. Sheng, "Report on wind turbine subsystem reliability survey of various databases," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/PR-5000-59111. [Online]. Available: http://www.nrel.gov/docs/fy13osti/59111.pdf

[14] C. Li *et al.*, "A comparison of fuzzy clustering algorithms for bearing fault diagnosis," *J. Intell. Fuzzy Syst.*, vol. 34, no. 6, pp. 3565–3580, Jun. 2018.

[15] D. Wang, K.-L. Tsui, and Q. Miao, "Prognostics and health management: A review of vibration based bearing and gear health indicators," *IEEE Access*, vol. 6, pp. 665–676, Nov. 2017.

[16] H. Darong, K. Lanyan, M. Bo, Z. Ling, and S. Guoxi, "A new incipient fault diagnosis method combining improved RLS and LMD algorithm for rolling bearings with strong background noise," *IEEE Access*, vol. 6, pp. 26001–26010, Apr. 2018.

[17] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[18] A. Soualhi, H. Razik, G. Clerc, and D. D. Doan, "Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system," *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2864–2874, Jun. 2014.

[19] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2727–2736, Mar. 2018.

[20] P. Bangalore and L. B. Tjernberg, "An artificial neural network approach for early fault detection of gearbox bearings," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 980–987, Mar. 2015.

[21] C. Li, J. L. V. de Oliveira, M. C. Lozada, D. Cabrera, V. Sanchez, and G. Zurita, "A systematic review of fuzzy formalisms for bearing fault diagnosis," *IEEE Trans. Fuzzy Syst.*, to be published.

[22] Y. Liu, B. He, F. Liu, S. Lu, and Y. Zhao, "Feature fusion using kernel joint approximate diagonalization of eigen-matrices for rolling bearing fault identification," *J. Sound Vibrat.*, vol. 385, pp. 389–401, Dec. 2016.

[23] M. H. C. Law and A. K. Jain, "Incremental nonlinear dimensionality reduction by manifold learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 377–391, Mar. 2006.

[24] L. Shi, L. Zhang, L. Zhao, L. Zhang, P. Li, and D. Wu, "Adaptive Laplacian eigenmap-based dimension reduction for ocean target discrimination," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 902–906, Jul. 2016.

[25] T. Hastie and W. Stuetzle, "Principal curves," *J. Amer. Statist. Assoc.*, vol. 84, no. 406, pp. 502–516, Jun. 1989.

[26] O. Schutze, X. Esquivel, A. Lara, and C. A. C. Coello, "Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 4, pp. 504–522, Aug. 2012.

[27] J. H. Ma, "Research on incipient fault fusion diagnosis methods for rotating machinery based on manifold learning," Ph.D. thesis, College Mech. Eng., Chongqing Univ., Chongqing, China, 2015.

[28] K. Levin and V. Lyzinski, "Laplacian eigenmaps from sparse, noisy similarity measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 1988–2003, Apr. 2017.

[29] Z. Zhang, J. Wang, and H. Zha, "Adaptive manifold learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 253–265, Feb. 2012.

[30] U. Ozertem and D. Erdogmus, "Principal curve time warping," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2041–2049, Jun. 2009.

[31] J. J. Verbeek, N. Vlassis, and B. Krose, "A k-segments algorithm for finding principal curves," *Pattern Recognit. Lett.*, vol. 23, pp. 1009–1017, Jun. 2002.

[32] Y. Gao, M. Wang, R. Ji, X. Wu, and Q. Dai, "3-D object retrieval with Hausdorff distance learning," *IEEE Trans. Ind. Electron.*, vol. 61, no. 4, pp. 2088–2098, May 2014.

[33] H. Cao, K. Zhou, X. Chen, and X. Zhang, "Early chatter detection in end milling based on multi-feature fusion and 3σ criterion," *Int. J. Adv. Manuf. Technol.*, vol. 92, nos. 9–12, pp. 4387–4397, Oct. 2017.

[34] J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services, "Bearing data set," NASA Ames Prognostics Data Repository, IMS, Univ. Cincinnati, Cincinnati, OH, USA, Tech. Rep., 2007. [Online]. Available: https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/

HONGJI REN received the B.S. degree in mechatronics engineering and the M.S. degree from Chongqing University, Chongqing, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree.

His current research interests include intelligent test and instruments, nondestructive testing and evaluation, and fault detection and diagnosis.

AIJUN YIN (M'13) received the B.S. degree in mechatronics engineering and the M.S. and Ph.D. degrees from Chongqing University, Chongqing, China, in 2001, 2003, and 2006, respectively.

He is currently a Professor with the College of Mechanical Engineering, Chongqing University. His current research interests include machine vision and image processing, machine learning, intelligent test and instruments, nondestructive testing and evaluation, modern signal analysis and processing, and fault detection and diagnosis.

QUAN ZHOU received the B.S. degree in mechatronics engineering and the M.S. degree from Chongqing University, Chongqing, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with The University of Tennessee, Knoxville.

His current research interests include computer vision, machine learning, pattern recognition, and image processing.

JIANG LI received the B.S. degree in mechanical design, manufacturing and automation from Huazhong Agricultural University, Wuhan, China, in 2014, and the M.S. degree from Chongqing University, Chongqing, China, in 2017.

His current research interests include machine learning, modern signal analysis and processing, and fault detection and diagnosis.

YIHUA HU (M'13–SM'15) received the B.S. degree in electrical motor drives and the Ph.D. degree in power electronics and drives from the China University of Mining and Technology, China, in 2003 and 2011, respectively.

Between 2011 and 2013, he was a Postdoctoral Fellow with the College of Electrical Engineering, Zhejiang University. From 2012 to 2013, he was an Academic Visiting Scholar with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne, U.K. Between 2013 and 2015, he was a Research Associate with the Power Electronics and Motor Drive Group, University of Strathclyde. He is currently a Lecturer with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, U.K. He has published more than 36 peer-reviewed technical papers in leading journals. His research interests include PV generation systems, power electronics converters and control, and electrical motor drives.