Research Article

An Approach to Fault Diagnosis for Gearbox Based on Image Processing

Yang Wang^{1,2} and Yujie Cheng^{1,2}

¹*School of Reliability and Systems Engineering, Beihang University, Beijing, China*

²*Science & Technology on Reliability and Environmental Engineering Laboratory, Beijing, China*

Correspondence should be addressed to Yujie Cheng; chengyujie@buaa.edu.cn

Received 2 June 2016; Accepted 3 August 2016

Academic Editor: Minvydas Ragulskis

Copyright © 2016 Y. Wang and Y. Cheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The gearbox is one of the most important parts of mechanical equipment and plays a significant role in many industrial applications. A fault diagnostic of rotating machinery has attracted attention for its significance in preventing catastrophic accidents and beneficially guaranteeing sufficient maintenance. In recent years, fault diagnosis has developed in the direction of multidisciplinary integration. This work addresses a fault diagnosis method based on an image processing method for a gearbox, which overcomes the limitations of manual feature selection. Differing from the analysis method in a one-dimensional space, the computing method in the field of image processing in a 2-dimensional space is applied to accomplish autoextraction and fault diagnosis of a gearbox. The image-processing-based diagnostic flow consists of the following steps: first, the vibration signal after noise reduction by wavelet denoising and signal demodulation by Hilbert transform is transformed into an image by bispectrum analysis. Then, speeded up robustness feature (SURF) is applied to automatically extract the image feature points of the bispectrum contour map, and the feature dimension is reduced by principal component analysis (PCA). Finally, an extreme learning machine (ELM) is introduced to identify the fault types of the gearbox. From the experimental results, the proposed method appears to be able to accurately diagnose and identify different types of faults of the gearbox.

1. Introduction

In modern industrial production, a gearbox is one of the most important transmission components of mechanical equipment because of its large power transmission capacity in a compact structure. However, the gearbox is acted upon by the impulse load, environmental corrosion, and fluctuating circulating stress in the work that is performed, which leads to a high failure rate. As an effective component of condition-based maintenance [1], the fault diagnosis has become prominent to guarantee the safe operation of gearboxes.

Gearbox conditions can be reflected by measurements of vibratory [2], acoustic [3], thermal [4], electrical [5], and oil-based signals [6]. Vibration signals are extremely sensitive to the existence of a gearbox, and a fault diagnosis based on the vibration signal is one of the most widely used methods. Generally, gearboxes are operated within harsh environments with severe noises and interferences [7]. The signal analysis

and fault characteristic parameters are rather complicated in the fault diagnosis of the gearbox. Generally, the field staff do not have enough professional knowledge to recognize the spectrum of the vibration signal and do not understand the principle of time domain frequency domain analysis. Therefore, it is difficult to diagnose the fault of the gearbox effectively. If an intelligent diagnosis of the fault of the gearbox can be achieved without extensive professional knowledge and feature extraction can be performed automatically, then the performance and efficiency of the diagnostic method will be greatly improved. Currently, fault diagnosis methods based on multiple disciplines have become the leading direction of development in the field of fault diagnosis. For example, a genetic algorithm of natural evolution theory (genetic algorithm), machine learning (support vector for the machine) [8], and bionics (ant colony algorithm) [9] have been applied to fault diagnosis in rotating machinery.

However, the calculation method of image processing has scarcely been applied to fault diagnosis for gearboxes. Fault diagnosis for a gearbox is essentially a process of fault pattern recognition, which is the same as the image classification process, belonging to the category of pattern recognition. Because the application of image feature extraction technology has been successful, this technology has high feasibility and a probable reference value for introducing the calculation method of image processing to the field of fault diagnosis.

In this paper, we present a novel approach that applies image feature extraction techniques to gearbox fault diagnosis. The method can accomplish a fully automatic feature extraction procedure for gearboxes with high accuracy and strong robustness.

The image expression of a vibration signal is a key step in the proposed method. A high-order spectrum method is one of the modern signal processing methods developed in recent years, and the high-order spectrum method plays an important role in non-Gaussian, nonlinear, noncausal, nonminimum phase and nonstationary signal processing [10, 11]. The bispectrum [12–14] is a subset of the higher order spectrum. Bispectrum preserves the phase information of the signals and can theoretically restrain the Gaussian noise. To avoid the interference of the working noise, a wavelet transform is applied to reduce the noise of the original vibration signal. Simultaneously, there are modulation components in the vibration signal spectrum, which cause difficulties in the identification of the characteristic frequency of the fault. It is therefore essential to demodulate the signal. This paper proposes an image conversion method based on bispectrum and Hilbert transform, and the images generated are used as inputs for feature extraction techniques.

In the development of the image automatic feature extraction technique in recent decades, the scale invariant feature transform (SIFT) method has become one of the most widespread image processing methods, with good robustness and high accuracy [15, 16]. However, the shortcomings of high resource consumption, high time complexity, and large computational time requirements constitute major limitations of SIFT. In 2006, Bay et al. proposed speeded up robust features (SURF) [17], which not only maintain the advantages of the high accuracy of the scale invariant feature transform (SIFT) algorithm but also overcome the shortcomings of its slow speed. Herein, SURF is employed to extract the feature points of the bispectrum contour map. The feature points extracted by SURF are described in terms of the 64-dimensional descriptor. To avoid the waste of resources for subsequent calculations, PCA is employed to reduce the dimension of the feature points.

The extreme learning machine (ELM), as an intelligent technology, has shown good performance in regression applications as well as in large datasets and multilabel classification applications [18]. Moreover, ELM has been proven to require less human intervention and less running time than most other pattern recognition methods. In this study, ELM was introduced to accomplish the state classification of the gearbox.

This paper is organized as follows: Section 2 introduces the related algorithms, Section 3 describes the case study

performed to validate the method, and Section 4 presents the conclusions and related future work.

2. Methodology

The procedure contains three major steps as follows: (1) image transformation of the vibration signal based on the bispectrum and Hilbert transform, (2) feature point extraction based on SURF-PCA, and (3) fault diagnosis based on ELM as illustrated in Figure 1.

2.1. Bispectrum Analysis Based on the Wavelet Transform Domain. Wavelet transform can be utilized to analyze a non-stationary signal with obvious time-frequency localization and multiresolution analysis ability, which can effectively enhance the transient information hidden in the mechanical signal. A bispectrum analysis is a powerful tool to analyze a non-Gaussian signal. A bispectrum analysis can characterize random signals from a higher probability structure, and, theoretically, it can suppress Gaussian noise completely. In a high-order spectrum, a bispectrum analysis has the minimal order but with all of the characteristics of a high-order spectrum. The analysis procedure of the bispectrum analysis based on the wavelet transform domain is shown in Figure 2.

In Figure 2, $n(t)$ is the interfering noise, which can be a Gaussian or a non-Gaussian noise. When $n(t)$ is a Gaussian noise, we can further process the signal without wavelet denoising. However, when $n(t)$ is a non-Gaussian noise, a bispectrum analysis is powerless to analyze the signal, and spectrum characteristics may be covered by the non-Gaussian noise. To eliminate non-Gaussian noise disturbance, the wavelet transform is employed to remove the noise. Some important parts of the gearbox, such as the gear and bearing, are typical rotating components. When the centralized or distributed fault occurred on these components, the vibration signal always shows strong nonstationary characteristics. However, these nonstationary vibration signals often show a strong characteristic of modulation. Therefore, it is necessary to demodulate the signal before the bispectrum analysis. The Hilbert transform is one of the most widely used demodulation methods [19]. Zero mean normalization is applied to the signal, denoising the signal, and then Hilbert transform is implemented on the signal. Finally, a bispectrum analysis is utilized to accomplish the conversion between the analytical signal and the image. The estimated method for a bispectrum analysis includes a parametric model and a nonparametric model. Compared with the indirect method, less computation is required by the direct method. Therefore, the direct estimation method is adopted in this paper. The flow of the algorithm is described as follows:

(1) By assuming that the observation data are finite in length, the sampling frequency is f_s . In the bispectrum domain, the number of the points is ω_1 and ω_2 ; so, the frequency sampling interval is $\Delta_0 = f_s/N_0$. Dividing $x(t)$ into the k th segment, each segment contains M points; that is, $N = KM$; then, subtract the mean of each sample.

(2) The discrete Fourier transform (DFT) is undertaken for the j th data segment; that is,

$$X^j(\lambda) = \frac{1}{M} \sum_{i=1}^M x^j(i) \exp\left(-j \frac{2\pi}{M} i \lambda\right). \quad (1)$$

FIGURE 1: Scheme for fault diagnosis.

FIGURE 2: Bispectrum analysis process in wavelet transform.

(3) According to the coefficient of the DFT, calculate the bispectrum estimation of each segment:

$$B_x^j(\lambda_1, \lambda_2) = \frac{1}{\Delta_0^2} \sum_{k_1=-L_1}^L \sum_{k_2=-L_1}^L X^j(\lambda_1 + k_1) X^j(\lambda_2 + k_2) \cdot X^j(\lambda_1 + k_1 + \lambda_2 + k_2). \quad (2)$$

(4) The bispectrum estimation is the mean value of K segment data; that is,

$$B_x(\omega_1, \omega_2) = \frac{1}{K} \sum_{j=1}^K B_x^j(\omega_1, \omega_2), \quad (3)$$

where $\omega_1 = [2\pi f_s/N_0]\lambda_1$ and $\omega_2 = [2\pi f_s/N_0]\lambda_2$.

2.2. SURF Descriptor

2.2.1. Interest Point Detection. A Hessian matrix is utilized to detect interest points by SURF, and the use of an integral image can greatly reduce the amount of calculation.

(1) *Integral Image.* The integral image can greatly increase the efficiency of box-type convolution filters. For a point X , $X = (x, y)$ in the integral image. The value of the point can be described as follows: the sum of all pixel values in a region formed by the origin point and point X is as follows:

$$I_\Sigma(X) = \sum_{i=0}^x \sum_{j=0}^y I(i, j). \quad (4)$$

Once an image is converted to an integral image, the sum of the gray levels in a rectangular region can be

FIGURE 3: (a) x direction, (b) y direction, and (c) xy direction.

calculated by 3 plus-minus calculations. We can see in the following discussion that the convolution templates used by SURF are all box templates, which can greatly improve the computational efficiency.

(2) *Approximate Hessian Matrix: H_{approx}* . The interest point detection of SURF is based on the Hessian matrix, which relies on the local maximum of the Hessian matrix determinant. When the Hessian matrix determinant appears to have a local maximum, the detection result is a blob-like structure, that is, a region brighter or darker than the surrounding area. Here, the register refers to the point and its small neighborhood. For a point $X = (x, y)$, the Hessian matrix is defined as

$$H(x, y) = \begin{bmatrix} L_{xx}(x, y) & L_{xy}(x, y) \\ L_{xy}(x, y) & L_{yy}(x, y) \end{bmatrix}, \quad (5)$$

where $L_{xx}(x, y)$, $L_{xy}(x, y)$, and $L_{yy}(x, y)$ are second-order partial derivatives and the two-dimension convolution of point (x, y) of the image I , respectively.

Instead of using a second-order Gaussian filter, a block filter is employed to approximate the two-order partial derivative of the Gaussian function to construct a fast-Hessian matrix. The approximate convolution template is utilized to process the integral image, with obvious benefits. The template is composed of simple rectangles, with computation independent of the size of the template. A 9×9 box filter is shown in Figure 3. In the original image, an image pyramid of a different size is formed by expanding the size of the box filter, and an integral image is applied to speed up the image convolution. Therefore, the approximate Hessian matrix can be obtained as

$$|H| = D_{xx}(X) D_{yy}(X) - [0.9 D_{xy}(X)]^2, \quad (6)$$

where D_{xx} , D_{yy} , and D_{xy} are the computational results of the box filter and the integral image. The extreme points are determined by the determinant and the eigenvalue of the matrix. If the determinant of H_{approx} matrix is positive and the two eigenvalues are different, the point is denoted as an extreme point.

According to the fast-Hessian matrix, the extremum of (x, y, σ) in the scale image can be obtained. First, nonmaximum suppression is carried out in the $3 \times 3 \times 3$ stereo neighborhood. Only the biggest of all points smaller than the

adjacent 26 points can be selected as the interest point. To locate (subpixel location) the candidate feature points, interpolation can be applied in a scale space and image space, and stable feature points and the scale value can be obtained [20].

2.2.2. Interest Point Location. To keep the characteristics of the rotation invariance of the interest point descriptor, the direction of the interest point should be determined first. Next, construct a Harr wavelet response in the direction of x -axis and y -axis, which is centered on an interest point, covering a 6σ radius. Give this response a different Gaussian weight coefficient; the closer the distance between feature points is, the larger the weight coefficient is. The sum of the Harr wavelet response is in the direction of x -axis and y -axis within 60° to form a local direction vector. Traverse the entire circular area, and the direction of the longest vector is selected as the principal direction of the interest point as shown in Figure 4.

Construct a window region with the center as a feature point and a side length of 20σ . Then, divide the window region into 4×4 subregions. The 5×5 sampling point is obtained from the subregion. Seek the wavelet response in the direction of x -axis and y -axis of each sampling point, designated as dx and dy , respectively. The Gaussian filter is performed as dx and dy in each subregion. The filter center is seen as the feature point, and a four-dimensional eigenvector $(\sum dx, \sum dy, \sum |dx|, \sum |dy|)$ is formed by summing dx , dy , $|dx|$, and $|dy|$ in each subregion. The four-dimensional eigenvector makes the 4 dimensions of the descriptor. Each descriptor consists of 4 dimensions. Thus, the $4 \times 4 \times 4 = 64$ dimensions are obtained, which are the SURF descriptor.

2.3. Extreme Learning Machine (ELM). ELM, proposed by Huang et al., was originally developed for single-hidden-layer feedforward neural networks and then extended to “generalized” single-hidden-layer feedforward networks (SLFNs). ELM is a novel learning algorithm with a faster learning speed and better generalization performance [21, 22]. The details of the ELM algorithm can be found in [18, 23–25].

The architecture of an SLFN can be described as shown in Figure 5, where x_i is the input sample, $\omega_i = [\omega_{1i}, \omega_{2i}, \dots, \omega_{ni}]$ represents the vector of link weights between all nodes in the input layer, $g(x)$ represents the activation function of neurons in the hidden layer, b_i represents the threshold of neurons in

FIGURE 4: Determine the main direction of the feature point.

FIGURE 5: Architecture of an SLFN.

the hidden layer, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ represents the vector of link weights between the i th node in the hidden layer and all nodes in the output layer, y_j represents the outputs of the network, n represents the number of nodes in the input layer, L represents the number of nodes in the hidden layer, m represents the number of nodes in the output layer, and $j = 1, 2, 3, \dots, m$. Then, the mathematical model of ELM can be described as

$$y_i = \sum_{j=1}^L \beta_j g_i(x_j) = \sum_{j=1}^L \beta_j g(\omega_j \cdot x_j + b_j). \quad (7)$$

Compared with SVM, whose kernel function greatly affects the results [21], ELM is not as sensitive to the activation function $g(x)$ as SVM, and almost all nonlinear piecewise continuous functions satisfying the ELM universal approximation capability theorems can be selected as the activation function [26], such as the sigmoid function, hard-limit function, or multiquadric function. The sigmoid function is the major activation function applied to the feedforward neural networks, and ELM with the hard-limit and multiquadric functions also shows good performance [23]. Thus, in this study, $g(x)$ is selected as a sigmoid function, and the process of ELM can be concluded as follows:

- (1) Determine the number, L , of neurons in the hidden layer and the activation function $g(x)$ and randomly assign ω_i , b_i , and β_i .

- (2) Calculate the output vector of the hidden layer.

- (3) Calculate the output weight $\hat{\beta}$.

3. Case Study

3.1. Experimental Facilities. The experimental data were obtained from the 2009 PHM data challenge competition. Data were collected from a two-class standard cylinder spur gear reducer. The reducer contains an input shaft, idler shaft, and an output shaft. The first stage reduction gear ratio is 1.5, and the second stage reduction gear ratio is 1.667. There are 32 teeth in the input shaft and 80 teeth in the output shaft. The two gears on the idler shaft both have 48 teeth. Figure 6 is a schematic of the gearbox used to collect the data and Figure 7 is the physical map of two-stage reducer.

The data were acquired using input shaft speeds of 35 Hz, 45 Hz, and 50 Hz. The sampling frequency is 66.7 Hz, and the sampling time is set to 66.7 kHz. The fault was detected as shown in Table 1.

3.2. Feature Extraction Based on Bispectrum and SURF. In this section, bispectrum and SURF were utilized to extract the feature vector, which is employed to process the two-class gear reducer signal. Considering the influence of environmental noise, wavelet denoising was performed on the gearbox signal, and then the signal was demodulated by Hilbert transform to obtain the analytic signal. Then,

FIGURE 6: Schematic of the apparatus.

TABLE 1: Fault information of gearbox.

		Normal	Fault 1	Fault 2	Fault 3	Fault 4
Gear	32T	Good	Chipped	Good	Good	Good
	96T	Good	Good	Good	Good	Good
	48T	Good	Eccentric	Eccentric	Good	Good
	80T	Good	Good	Good	Broken	Good
Bearing	IS:IS	Good	Good	Good	Inner	Inner
	ID:IS	Good	Good	Good	Ball	Good
	OS:IS	Good	Good	Good	Outer	Good
	IS:OS	Good	Good	Good	Good	Good
	ID:OS	Good	Good	Good	Good	Good
	OS:OS	Good	Good	Good	Good	Good
Shaft	Input	Good	Good	Good	Imbalance	Good
	Output	Good	Good	Good	Good	Keyway sheared

FIGURE 7: Physical map of two-stage reducer.

bispectrum was employed to analyze the analytic signal. Herein, the direct method was adopted, and the bispectrum contour map of different fault modes is shown below.

Four fault modes were collected in this section, namely, fault 1, fault 2, fault 3, and normal condition. Specific information is shown in Table 1. The length of the data is 5000 in each group. Figures 8–12 are bispectrum contour maps of faults

1–4, respectively. Three datasets have been selected for the comparison. From the contour map, we can see that different failure modes have different frequency distributions, which can reflect the differences between the different fault modes.

In the experiment, the SURF descriptor extracted feature points with 64 dimensions. Then, principal component analysis (PCA) was utilized to reduce the dimensions to 20 dimensions. The first three principal components are presented in a three-dimensional space as shown in Figure 13, from which we can conclude that the features separately show a strong ability. However, to identify the failure modes accurately, the feature must be combined with a classifier.

(3) *Fault Diagnosis Results Based on ELM.* The input eigenvector of ELM was extracted by bispectrum and SURF. For each fault mode, 240 sets of data were collected. The 80 sets were selected randomly for training the extreme learning machine. The other 160 sets were selected as test data. Usually, fault diagnosis accuracy is selected as the metric of the effectiveness of fault diagnosis method, as in [7, 27, 28]. In this study, fault diagnosis accuracy is also selected to verify the proposed method. To display the results visually, the

FIGURE 8: Bispectrum contour map of normal condition.

FIGURE 9: Bispectrum contour map of fault 1.

FIGURE 10: Bispectrum contour map of fault 2.

classification result was contrasted with the actual result. The accuracy rate is defined as the odds ratio of the right result and the total result. The diagnosis result is shown in Figure 14(a).

For comparison, the proposed method is compared with traditional diagnosis method based on empirical mode decomposition (EMD). EMD, as an effective decomposition method for nonlinear and nonstationary signals, has been applied to fault diagnosis of gearbox. However, the problem of mode mixing which existed in the EMD makes Figures 14(a) and 14(b) unable to fully decompose a signal with intermittent components, which will affect the accuracy of fault diagnosis. Herein, EMD is used to decompose the origin signal into several intrinsic mode functions (IMFs) and the singular value is used as feature vector. Then put them into the ELM classifier. The fault diagnosis result is shown in Figure 14(b).

In Figure 14 the horizontal ordinate represents the normal condition and 4 fault conditions, while the vertical ordinate represents the numbers of correct diagnosis classifications. The diagnosis accuracy of each fault mode using both diagnosis method based on image processing and traditional EMD is listed in Table 2 in detail.

From Table 2, we can see that, in the proposed fault diagnosis method based on image processing, the diagnosis accuracy of normal condition and 4 fault conditions reaches 96.25%, 98.75%, 97.5%, 87.50%, and 97.5%, respectively, among which the highest diagnosis accuracy reaches 98.75%, while the lowest is 87.5%. The average diagnosis accuracy of the five conditions is 95.5%, indicating the high effectiveness of the proposed method based on image processing.

The diagnosis accuracy using traditional method based on EMD under each fault mode is 90.625%, 92.5%, 91.25%, 79.375%, and 91.25%, respectively. The highest accuracy, 92.5%, is 6.25% lower than the proposed diagnosis method based on image processing. While the lowest accuracy, 79.375%, is 8.125% lower than the proposed method. The average accuracy of the five fault modes is 89%, 6.5% lower than the proposed method.

The diagnosis accuracy results are listed in Figure 15 in histogram for a clearer comparison. From Figure 15, it can be seen that, for each fault mode, the fault diagnosis based on bispectrum and SURF has higher accuracy, demonstrating the superiority of the proposed method.

4. Conclusions

In this paper, a novel fault diagnosis method for a gearbox based on the image processing technology is first proposed. The method contains the following three main steps. (1) The first step is image generation process: a bispectrum analysis is applied to transform the vibration after wavelet denoising and Hilbert demodulation. (2) The second step is feature

TABLE 2: Fault diagnosis result.

Fault types	Testing samples	Correct diagnosis results		Accuracy	
		Bispectrum and SURF	EMD	Bispectrum and SURF	EMD
Normal	160	154	145	96.25%	90.625%
Fault 1	160	158	148	98.75%	92.50%
Fault 2	160	156	146	97.50%	91.25%
Fault 3	160	140	127	87.50%	79.375%
Fault 4	160	156	146	97.50%	91.25%
Total	800	764	712	95.5%	89.00%

FIGURE 11: Bispectrum contour map of fault 3.

FIGURE 12: Bispectrum contour map of fault 4.

FIGURE 13: The first three features extracted using PCA.

extraction process: SURF is introduced to extract the feature points of the bispectrum contour map automatically. (3) The third step is fault pattern recognition process: based on the features extracted by SURF, ELM is then used to identify the fault types of the gearbox. Differing from the traditional fault diagnosis method of vibration analysis in a one-dimensional space, the method applying the computing method to a field of image processing in a 2-dimensional space accomplished autoextraction and fault diagnosis of the gearbox. Almost no manual intervention was needed for the whole diagnosis process, which avoids the limitations of other methods that require a large amount of expert knowledge. Simultaneously, the results of the case study fully demonstrated the effectiveness and high diagnostic accuracy of the image-processing-based method for the gearbox.

Our subsequent work will be focused on the following:

- (1) In view of the complex structure of the gearbox, more fault modes should be taken into account. Therefore, we should apply this method to more fault types of the gearbox.

FIGURE 14: (a) Fault diagnosis result of ELM based on EURF. (b) Fault diagnosis result of ELM based on EMD.

FIGURE 15: Result comparison of EMD and SURF.

- (2) We should improve the computing speed while maintaining diagnostic accuracy.

Competing Interests

The authors declare that there are no potential competing interests in this research.

Acknowledgments

This study was supported by the Fundamental Research Funds for the Central Universities (Grant no. YWF-16-BJ-J-18), the National Natural Science Foundation of China (Grant no. 51575021), and the Technology Foundation Program of National Defense (Grant no. Z132013B002).

References

- [1] D. Galar, A. Thaduri, M. Catelani, and L. Ciani, "Context awareness for maintenance decision making: a diagnosis and prognosis approach," *Measurement*, vol. 67, pp. 137–150, 2015.
- [2] Y. Lei, N. Li, J. Lin, and Z. He, "Two new features for condition monitoring and fault diagnosis of planetary gearboxes," *JVC/Journal of Vibration and Control*, vol. 21, no. 4, pp. 755–764, 2015.
- [3] J. Hou, W. Jiang, and W. Lu, "Application of a near-field acoustic holography-based diagnosis technique in gearbox fault diagnosis," *Journal of Vibration and Control*, vol. 19, no. 1, pp. 3–13, 2013.
- [4] A. M. D. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating machinery using infrared thermal image," *Expert Systems with Applications*, vol. 39, no. 2, pp. 2082–2091, 2012.
- [5] J. R. Ottewill and M. Orkisz, "Condition monitoring of gearboxes using synchronously averaged electric motor signals," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 482–498, 2013.
- [6] C. Li and M. Liang, "Extraction of oil debris signature using integral enhanced empirical mode decomposition and correlated reconstruction," *Measurement Science and Technology*, vol. 22, no. 8, Article ID 085701, 2011.
- [7] C. Li, R. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals," *Mechanical Systems and Signal Processing*, vol. 76–77, pp. 283–293, 2016.
- [8] L. Lu, J. Yan, and C. W. De Silva, "Dominant feature selection for the fault diagnosis of rotary machines using modified genetic algorithm and empirical mode decomposition," *Journal of Sound and Vibration*, vol. 344, pp. 464–483, 2015.
- [9] X. Zhang, W. Chen, B. Wang, and X. Chen, "Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization," *Neurocomputing*, vol. 167, pp. 260–279, 2015.
- [10] K. M. Hock, "Narrowband weak signal detection by higher order spectrum," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 874–879, 1996.

- [11] X. Zhang, *Time Series Analysis-Higher Order Statistics Method*, vol. 1, Tsinghua University Press, Beijing, China, 1996.
- [12] J. Tretrong, J. K. Sinha, F. Gu, and A. Ball, "Bispectrum of stator phase current for fault detection of induction motor," *ISA Transactions*, vol. 48, no. 3, pp. 378–382, 2009.
- [13] F. E. H. Montero and O. C. Medina, "The application of bispectrum on diagnosis of rolling element bearings: a theoretical approach," *Mechanical Systems and Signal Processing*, vol. 22, no. 3, pp. 588–596, 2008.
- [14] L. Saidi, F. Fnaiech, H. Henao, G.-A. Capolino, and G. Cirrincione, "Diagnosis of broken-bars fault in induction machines using higher order spectral analysis," *ISA Transactions*, vol. 52, no. 1, pp. 140–148, 2013.
- [15] S. Yue, Q. Wang, and R. Zhao, "Robust wide baseline feature point matching based on scale invariant feature descriptor," in *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues*, pp. 329–336, Springer, 2008.
- [16] Y. Wang, X. Ban, J. Chen, B. Hu, and X. Yang, "License plate recognition based on SIFT feature," *Optik—International Journal for Light and Electron Optics*, vol. 126, no. 21, pp. 2895–2901, 2015.
- [17] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: speeded up robust features," in *Computer Vision—ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds., vol. 3951 of *Lecture Notes in Computer Science*, pp. 404–417, Springer, Berlin, Germany, 2006.
- [18] G.-B. Huang, X. Ding, and H. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, no. 1–3, pp. 155–163, 2010.
- [19] W. Teng, X. Ding, X. Zhang, Y. Liu, and Z. Ma, "Multi-fault detection and failure analysis of wind turbine gearbox using complex wavelet transform," *Renewable Energy*, vol. 93, pp. 591–598, 2016.
- [20] J. Li, Y. Wang, and Y. Wang, "Visual tracking and learning using speeded up robust features," *Pattern Recognition Letters*, vol. 33, no. 16, pp. 2094–2101, 2012.
- [21] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 985–990, July 2004.
- [22] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [23] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [24] G.-B. Huang, "An insight into extreme learning machines: random neurons, random features and kernels," *Cognitive Computation*, vol. 6, no. 3, pp. 376–390, 2014.
- [25] V. Malathi, N. S. Marimuthu, S. Baskar, and K. Ramar, "Application of extreme learning machine for series compensated transmission line protection," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 5, pp. 880–887, 2011.
- [26] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *IEEE Transactions on Neural Networks*, vol. 17, no. 4, pp. 879–892, 2006.
- [27] X. Jiang, S. Li, and Y. Wang, "A novel method for self-adaptive feature extraction using scaling crossover characteristics of signals and combining with LS-SVM for multi-fault diagnosis of gearbox," *Journal of Vibroengineering*, vol. 17, no. 4, pp. 1861–1878, 2015.
- [28] L. Zhao, W. Yu, and R. Yan, "Gearbox fault diagnosis using complementary ensemble empirical mode decomposition and permutation entropy," *Shock and Vibration*, vol. 2016, Article ID 3891429, 8 pages, 2016.