Springer

KSME

Journal of Mechanical Science and Technology 31 (2) (2017) 545~553
www.springerlink.com/content/1738-494x(Print)/1976-3824(Online)
DOI 10.1007/s12206-017-0107-3

Gear fault diagnosis under variable conditions with intrinsic time-scale decomposition-singular value decomposition and support vector machine †

Zhanqiang Xing[1,2], Jianfeng Qu[1,2,*], Yi Chai[1,2], Qiu Tang[1,2] and Yuming Zhou[1,2]

[1]*Key Laboratory of Dependable Service Computing in Cyber Physical Society of Ministry of Education, Chongqing University, Chongqing, 400030, China*
[2]*College of Automation, Chongqing University, Chongqing, 400030, China*

(Manuscript Received March 22, 2016; Revised August 18, 2016; Accepted September 30, 2016)

Abstract

The gear vibration signal is nonlinear and non-stationary, gear fault diagnosis under variable conditions has always been unsatisfactory. To solve this problem, an intelligent fault diagnosis method based on Intrinsic time-scale decomposition (ITD)-Singular value decomposition (SVD) and Support vector machine (SVM) is proposed in this paper. The ITD method is adopted to decompose the vibration signal of gearbox into several Proper rotation components (PRCs). Subsequently, the singular value decomposition is proposed to obtain the singular value vectors of the proper rotation components and improve the robustness of feature extraction under variable conditions. Finally, the Support vector machine is applied to classify the fault type of gear. According to the experimental results, the performance of ITD-SVD exceeds those of the time-frequency analysis methods with EMD and WPT combined with SVD for feature extraction, and the classifier of SVM outperforms those for K-nearest neighbors (K-NN) and Back propagation (BP). Moreover, the proposed approach can accurately diagnose and identify different fault types of gear under variable conditions.

Keywords: Gear fault diagnosis; Intrinsic time-scale decomposition; Singular value decomposition; Support vector machine

1. Introduction

As any defect of the gearbox can result in machine downtime and loss of production, fault diagnosis of gearbox plays an important role in industrial machinery [1]. Because of the complex structure and poor work conditions, the gear teeth may easily suffer damages such as wear, pitting and pitting wear on teeth surface [2]. Fault diagnosis mainly includes three parts: Feature extraction, pattern recognition, and classification [3]. Therefore, an optimal method for fault diagnosis is to find a feature extraction technique that can extract the most distinctive features to optimize classification results and improve diagnostic accuracy [4]. In gear fault diagnosis, the variation of working conditions often influences the generated vibration patterns. For this reason, it is very important to seek an efficient fault diagnosis method applicable to variable conditions.

Recently, the study of gear fault diagnosis has received extensive concern from the researchers. Raad et al. employed an indicator of cyclostationarity for gear diagnosis in time domain [5]. Pavle et al. successfully performed a fault detection of the mechanical drives based on the wavelet packet Renyi entropy signatures [6]. Liu et al. performed the gear fault detection by employing the fast dynamic time warping to identify the corresponding faulty gear with small fluctuations of the operating speed of the machine [7]. Li et al. addressed a multimodal deep support vector classification method based on deep learning strategy to perform fault diagnosis tasks for gearboxes, which was proven to be effective for the gearbox fault diagnosis [8]. Zhang et al. adopted the energy operator demodulating of optimal resonance components to detect the compound faults of gearboxes [9]. Through these studies, rich experience in gear fault diagnosis has been accumulated and the pivotal function under variable conditions. However, both the nonlinear and non-stationary characteristic of vibration signals and the interferences under variable conditions increase the difficulty of extracting features from the complex vibration signal, and the fault diagnosis under variable conditions still needs to be further studied.

At present, several popular time-frequency analysis methods for gear fault feature extraction have been suggested in the literature, such as the Short-time fourier transform (STFT) [10], the Wavelet packet transform (WPT) [11] and the Wigner-Ville distribution [12, 13]. However, those methods also have some weakness, for example the major limitation of SFTF is unavoidable in trade-off between time and frequency resolution, WPT is not a self-adaptive time-frequency analysis method and the

*Corresponding author. Tel.: +86 23 65106464, Fax.: +86 23 65112174
E-mail address: qujianfeng@cqu.edu.cn
† Recommended by Associate Editor Byeng Dong Youn

computation of WPT is time-consuming, the Wigner-Ville distribution is limited by the presence of cross-term interference. Hence, several self-adaptive signal processing methods have been proposed, such as Empirical mode decomposition (EMD) [14] and Local mean decomposition (LMD) [15]. Nonetheless, the EMD method also has some drawbacks such as over envelope, mode mixing [16] and end effects [17]. The LMD method still has some inevitable problems such as distorted components and time-consuming decomposition.

Intrinsic time-scale decomposition (ITD) method was proposed by Frei et al. [18] as a novel time frequency analysis method for analyzing the non-stationary and nonlinear signals. With high decomposition efficiency and frequency resolution, ITD can help decompose a complex signal into several PRCs with the lowest variation of PRCs as the trend, which can accurately extract the dynamic feature of nonlinear signals. Meanwhile, ITD does not involve spline interpolation and screening process. It has low edge effect and is promising for the application to the real-time data analysis [19]. Due to its advantages in adaptability and higher computing efficiency, the ITD method is well suitable for gear fault diagnosis under variable conditions. Lin et al. [20] performed a bearing fault diagnosis method based on spectral kurtosis and ITD which is able to improve the performance of the bearing fault diagnosis.

Under variable conditions, the fault time frequency characteristic of gear changes over time, and the ITD method is suitable for dealing with the complex signal. Nevertheless, the proper rotation components obtained through ITD under variable conditions are always too tremendous and complex to be regarded as the fault feature vectors. Thus, time-frequency analysis ITD is combined with singular value decomposition for the first time in this study for gear fault diagnosis. SVD is applied to achieve feature reduction and improve the robustness of the feature vectors. The singular value can maintain the stability and enhance the robustness of the feature extraction under variable conditions. Support vector machine (SVM) as a powerful machine learning method has been successfully applied to fault classification [21]. Moreover, SVM requires less human intervention and less running time than the other intelligent classification methods such as K-nearest neighbors (K-NN) [22] and Back propagation [23].

In this study, a novel hybrid model based on ITD, SVD and SVM is presented for gear fault diagnosis under variable conditions. Firstly, the vibration signal is decomposed into some PRCs with ITD to obtain the feature vectors under variable conditions. Secondly, SVD is applied to achieve feature reduction and improve the robustness of the feature vectors. The proposed method is compared with EMD-SVD and WPT-SVD in terms of feature extraction. Finally, the singular values are fed into the trained SVM to identify the fault type.

This paper is organized as follows: Sec. 2 presents the introduction of ITD, SVD and SVM, Sec. 3 shows the case study performed to validate the method, and Sec. 4 concludes the paper.

2. Methodology

Gear fault diagnosis includes fault feature extraction and fault classification. The vibration signal processing method is one of the most important fault feature extraction methods. According to the extracted features, the different fault types can thus be identified [24]. The detailed process will be introduced in the following part of the section.

2.1 Signal decomposition based on ITD

As a new self-adaptive signal decomposition method, Intrinsic Time-scale decomposition is applied to the non-stationary signal analysis successfully. Any complex non-stationary vibration signal can be decomposed into a series of PRCs and a monotonous trend signal [25]. For an original signal X_t, the operator L is defined as the baseline extraction factor, and the rest signal is regarded to have a proper rotation. The decomposition of the X_t can be represented as follow:

$$X_t = LX_t + (1-L)X_t = L_t + H_t \tag{1}$$

where $L_t = LX_t$ is the baseline signal, $H_t = (1-L)X_t$ is the proper rotation component.

It is assumed that $X_k, k = 1,2,...$ is the extremum of the signal X_t, τ_k is the corresponding moment, and $\tau_0 = 0$. In order to simplify the definition, X_k and L_k are defined as the $X(\tau_k)$ and $L(\tau_k)$, respectively. As L_t and H_t are defined on $[0, \tau_k]$, and X_t is also meaningful in $[0, \tau_{k+2}]$, the baseline extracting operator ε on interval $(\tau_k, \tau_{k+1}]$ can be expressed as follow:

$$LX_t = L_t = L_k + \frac{L_{k+1} - L_k}{X_{k+1} - X_k}(X_t - X_k), t \in (\tau_k, \tau_{k+1}]. \tag{2}$$

The L_{k+1} in Eq. (2) can be expressed as follow:

$$L_{k+1} = \alpha[X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}(X_{k+2} - X_k)] + (1-\alpha)X_{k+1} \tag{3}$$

where α is a linear gain, and $0 < \alpha < 1$. In general, $\alpha = 0.5$. The proper rotation component can be expressed as below:

$$HX_t = (1-L)X_t = H_t = X_t - L_t. \tag{4}$$

Base signal L_t retains the monotonicity between the extreme points of the signal. Besides, the local relatively high frequency is extracted as a proper rotation component. Then, the baseline signal is set as the input signal, and the above steps are repeated until a drab trend signal is obtained. Consequently, the original signal is decomposed into a set of proper rotation components with a frequency range is from high to low and a monotonic trend component. The overall procedure can be expressed with the formula below:

$$\begin{aligned} X_t &= HX_t + LX_t = HX_t + (H+L)LX_i = [H(1+L)+L^2]X_t \\ &= (H\sum_{k=0}^{p-1} L^k + L^p)X_t = H_t^1 + H_t^2 + \cdots + H_t^p + L_t^p \end{aligned} \tag{5}$$

wherein H_t^i is the i_{th} layer of the proper rotation, L_t^p is the monotonic baseline signal representing the trend of X_t, and p is the decomposition level.

2.2 Feature extraction based on SVD

Under variable conditions, the vibration signal of gear is decomposed with ITD into some PRCs which are too complex to be regarded as the fault feature vectors. The SVD method is proposed to reduce the dimensions of the fault feature vectors and improve the accuracy of the classification [26].

Singular value decomposition is an orthogonal matrix transformation algorithm, and the singular value has a good numerical stability when the matrix elements change. It is assumed that A is a $m \times n$ matrix. With the SVD method, any matrix $A \in R_{m*n}$ can be decomposed into three matrices, namely $U \in R_{m*m}, S \in R_{m*n}$ and $V \in R_{m*n}$ as shown below:

$$A_{m\times n} = U_{m\times m} S_{m\times n} (V_{n\times n})^T \tag{6}$$

wherein U and V are orthogonal, and S is a diagonal matrix with $r = rank(A)$, as shown below:

$$S = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{r-1} & \vdots \\ 0 & 0 & \cdots & 0 & \sigma_r \end{pmatrix}. \tag{7}$$

σ_i represents the singular values of A. The singular values are listed in descending order as follow:

$$\sigma_1(A) \ge \sigma_2(A) \ge \cdots \ge \sigma_r(A). \tag{8}$$

SVD is adopted to reduce the dimensionality of the PRCs obtained through ITD. The feature of the original matrix can be expressed with the singular values, which is beneficial to compressing the scale of the feature vector.

2.3 State classification based on SVM

The feature vectors of the different fault types are obtained based on ITD and SVD. In addition, the SVM is proposed to realize the classification of multiple classes. As a statistical learning theory proposed by Vapnik-Chervonenkis, SVM has a huge advantage in solving the small sample, nonlinear and high dimensional pattern recognition problem [27]. For two-class problem, a training set $(x_k, y_k), k = 1,2...N$ is given, where $x_k \in R^n$ and $y_k \in \{-1,1\}$. The mechanism of SVM is to construct an optimal hyperplane and a classifier of the form:

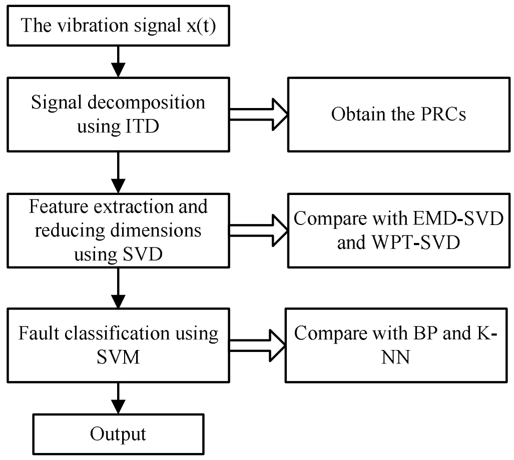

Fig. 1. Block diagram of the proposed method.

$$y(x) = sign[w^T x + b] \tag{9}$$

wherein w,b are utilized for defining a hyperplane in the feature space consisting of data vectors x that satisfy the constraint $w \cdot x + b = 0$. The classifier can be formulated as follow:

$$y_k[w^T x + b] \ge 1 - \xi_k, \xi_k \ge 0, k = 1,...,N. \tag{10}$$

The classification interval can be calculated as $2/\|w\|$, and the problem of constructing an optimal hyperplane is converted under the constraint of:

$$\min J_p(w,\xi) = \frac{1}{2} w^T w + c\sum_{k=1}^{N} \xi_k \tag{11}$$

wherein c is a positive real constant that controls the punishment for misclassified samples.

In order to solve the constrained optimization problem, the Lagrangian is introduced, the conditions for optimality are created, and finally the problem is solved in the dual space [28].

For multi-classification problems, several approaches to extend two-class SVMs to a multi-class SVM for multi-category classifications have been proposed [29]. The one-against-one algorithm that transforms a c-class problem into $c(c-1)/2$ two-class problems is applied in this work. Through voting, the class with the most votes is regarded as the class of the sample.

The proposed gear fault diagnosis technique under variable conditions is described in Fig. 1, with detailed steps shown as follows:

Step 1: Decompose the vibration signal into several PRCs and the trend signal by using ITD. Combine the PRCs and the trend signal into the feature matrix A.

Step 2: Compute the singular values of the feature matrix A with SVD. Regard the singular values as the feature vectors of the vibration signal.

Step 3: Diagnose the failure types according to the feature vectors by using SVM.

3. Experiments and analysis

3.1 Gear fault data

An experimental analysis on gear fault diagnosis is conducted to evaluate the effectiveness of the proposed method. The gear vibration data are obtained through rotating machinery vibration analysis with a QPZZ-II test platform, as shown in Fig. 2. The four types of the vibration signal (normal, pitting, wearing and pitting-wearing signal) are collected at 5120 sample/s. Under the working condition with no-load, the motor drives the input shaft at a motor speed of 880 RPM. In order to change the load, the current of the motor is controlled at 0.05 A, 0.1 A and 0.2 A (corresponding to the motor speeds of 866, 850 and 834 RPM, respectively). The length of each original collected signal is about 50000, and 25 samples for each vibration condition are extracted, as shown in Table 1. The vibration signal waveforms of different fault types under the no-load condition are shown in Fig. 3.

Table 1. Relevant information for the dataset.

Groups					
Label	Status	Working conditions			
		880 RPM	866 RPM	850 RPM	834 RPM
1	Normal	25	25	25	25
2	Pitting fault	25	25	25	25
3	Wearing fault	25	25	25	25
4	Pitting-wearing fault	25	25	25	25

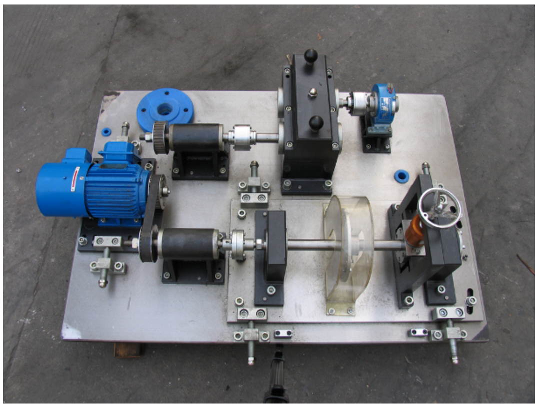

Fig. 2. Test platform of gearbox.

3.2 Gear fault diagnosis based on the proposed method

To acquire the fault feature vectors, the ITD method is used to decompose the vibration signals of the different fault types into the proper rotation components, as shown in Fig. 4. The original signals are decomposed into three PRCs and a monotonous trend signal. PCR1~PCR3 are the proper rotation components to represent the local relatively high frequency of the original signals and the L is the monotonous trend signal. The high frequency of the proper rotation components contains the primary information of the different fault types, and the difference among the fault types is significant. The PRCs and the trend signal is combined into a feature matrix, which can be constructed as fault feature vectors.

Under variable conditions, the PRCs obtained with ITD are too tremendous and complex to be regarded as the fault feature vectors. To solve this problem, the SVD is applied to reduce the dimensions of the fault feature vectors. After obtaining the feature matrix, the singular values can be obtained by conducting SVD. The singular value can maintain the stability and enhance the robustness of the fault diagnosis under variable conditions. The fault feature values obtained through ITD-SVD are partly shown in Table 2. The first three singular values are shown in Fig. 5. The three singular values represent the decomposed components of the vibration signal. It can be seen that the singular values under the same failure mode are

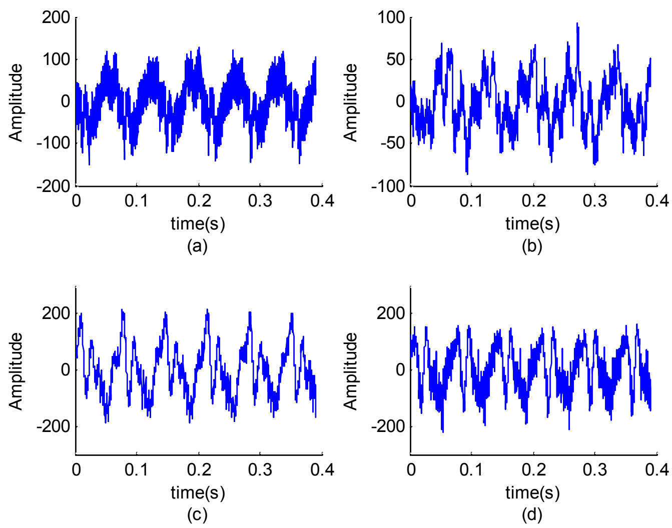

Fig. 3. Vibration signal waveforms of (a) the normal signal, and signals with (b) pitting fault; (c) wearing fault; (d) pitting-wearing fault under the no-load condition.

Table 2. Fault feature values obtained through ITD-SVD.

Condition	Fault feature values							
	1	2	3	4	5	6	7	8
Normal	1465.6	1465.1	1467.4	1454.7	1443.0	1451.2	1451.7	1456.7
	681.7	638.4	586.1	749.4	858.7	695.9	709.8	691.5
	397.5	404.9	363.1	358.7	360.7	261.6	341.2	349.2
Pitting fault	645.2	599.5	577.4	791.3	760.7	654.6	587.0	616.9
	463.7	463.6	460.8	458.3	468.2	446.0	455.1	456.0
	245.9	324.4	266.0	326.0	289.4	345.9	283.8	323.1
Wearing fault	2446.1	2346.0	2178.8	1764.5	1418.1	1683.7	1632.92	1585.8
	682.0	690.3	700.8	1029.1	879.8	1018.5	1001.5	1003.2
	606.0	611.5	631.1	645.4	659.3	718.1	740.0	658.5
Pitting-wearing fault	1829.9	1820.2	1534.5	1676.7	1669.0	1630.2	1873.1	1716.4
	1301.7	1314.5	1355.0	1273.1	1324.4	1316.8	1281.2	1356.3
	802.4	735.3	875.5	934.9	807.2	878.6	874.8	811.3

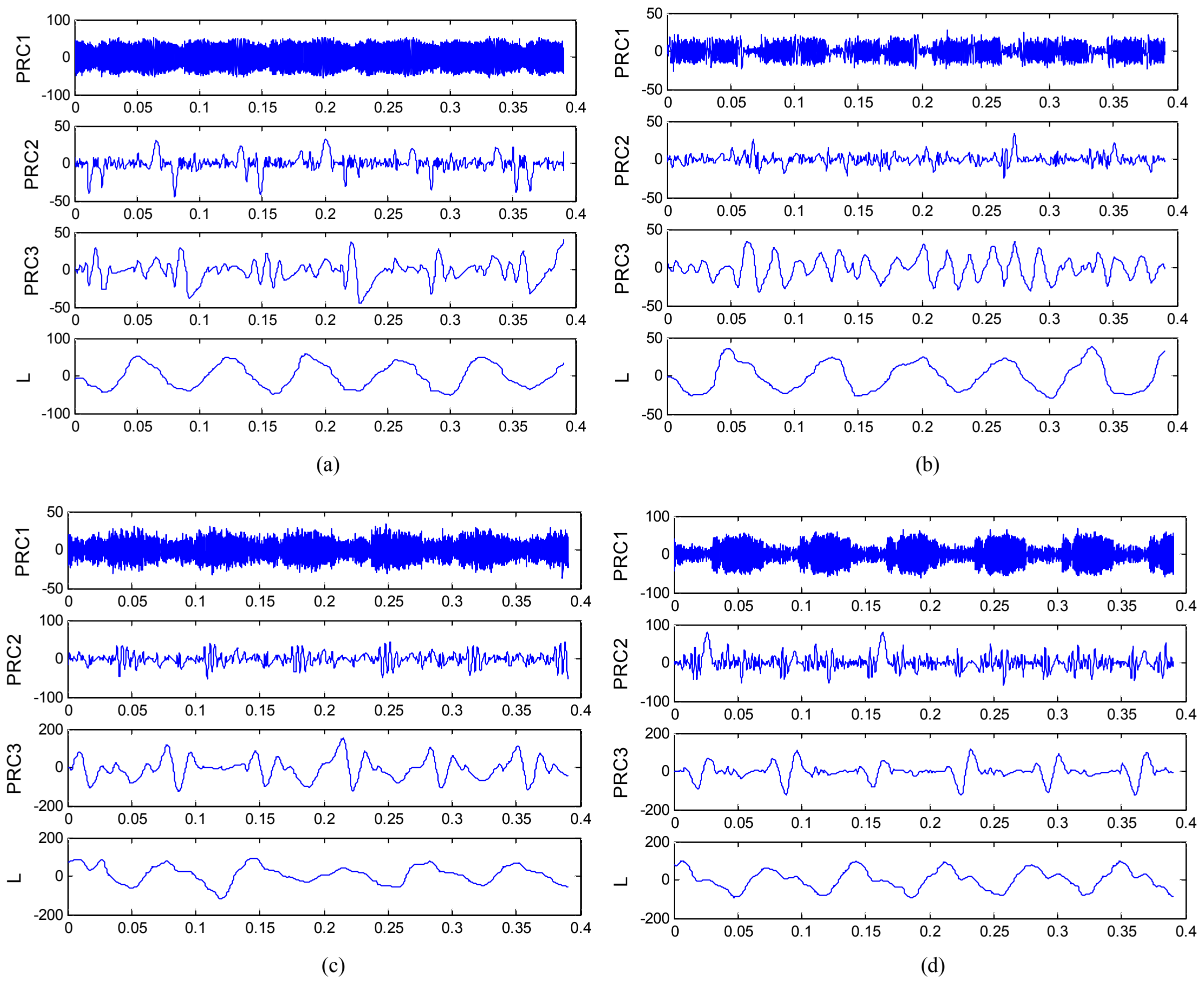

Fig. 4. PRCs of the vibration signal obtained with ITD: (a) Normal status; (b) pitting fault; (c) wearing fault; (d) pitting-wearing fault.

highly consistent and the difference of the singular values under different fault types is large enough to easily separate them. Due to the significant separability under variable conditions, the singular value can be applied to the fault classifier.

Through the proposed feature extraction method based on ITD and SVD, singular value vectors are obtained as the feature vectors. On the basis of SVM, state classification can be done for recognizing the fault types of gear under variable conditions. The datasets in Table 1 are divided into a training dataset with 20 samples of data selected randomly from each fault type and a testing dataset with 380 samples of data. The classification result of ITD-SVD is shown in Fig. 6, from

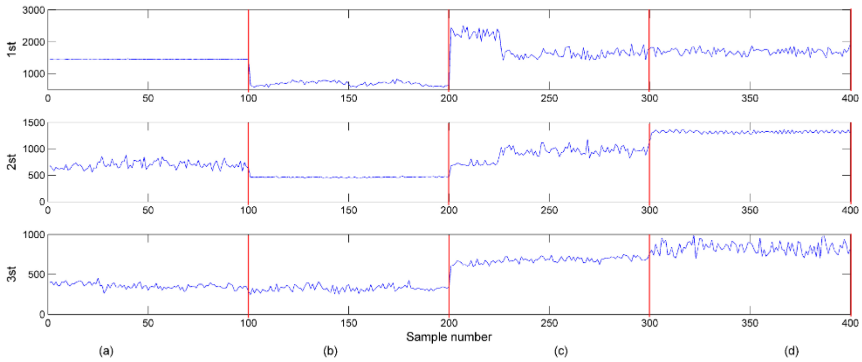

Fig. 5. The first three singular values obtained through ITD-SVD in the dataset: (a) Normal status; (b) pitting fault; (c) wearing fault; (d) pitting-wearing fault (Note: Sample number. 1-100, 101-200, 201-300, 301-400 represent normal status, pitting fault, wearing fault, pitting-wearing fault, respectively).

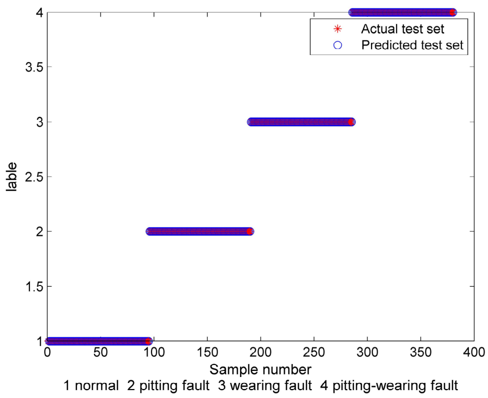

Fig. 6. Classification result of SVM.

which it can be seen that the actual output of SVM is extremely consistent with the target output under variable conditions. The average classification accuracy of the three feature extraction methods based on SVM is 99.53 %. In the case of small samples, the SVM can achieve a good classification result.

3.3 Comparisons with EMD-SVD and WPT-SVD

In this subsection, some conventional time-frequency analysis methods such as EMD and WPT combined with SVD are also applied to the feature extraction of gear under variable conditions, and their performances are compared with that of the proposed approach.

With the EMD method, the vibration signal can be decomposed into some IMFs and a residual. The SVD method is applied to the first 4 IMFs to obtain the feature vector, and the first three singular values for datasets in Table 1 are shown in Fig. 7. The classification interval between the normal signal and the signal with pitting fault is so prominent that the pitting fault can be identified easily. Nevertheless, the classification of the wearing and the pitting-wearing fault is mixed up completely, which may result in misclassification. In other words, the feature vector obtained by EMD and SVD is affected greatly by working condition variation.

With the WPT method, the vibration signal can be divided into 3 levels to obtain the wavelet packet coefficients of the 8 wavelet packet nodes using dB5. Furthermore, the wavelet packet coefficients are utilized to reconstruct the signals to acquire the decomposed signal, and the SVD method is applied to the decomposed signal to obtain the feature vector. The first three singular values are shown in Fig. 8. The wearing fault is divided into two parts. As a result, the WPT method combined with SVD is sensitive to gear fault diagnosis under variable conditions, meanwhile it is also not suited for gear fault diagnose under variable conditions.

Through the contrast analysis, it is proved that the feature extraction method based on ITD-SVD can maintain the stability and enhance the robustness of the gear fault diagnosis under variable conditions. The singular values under variable conditions are shown in Fig. 9. For gear feature extraction under variable conditions, the proposed method can obviously improve the performance of fault patterns and outperforms the time-frequency analysis methods: EMD and WPT combined with SVD.

The average classification accuracy of the three feature extraction methods based on SVM is shown in Table 3. It is verified again that the proposed method outperforms the time-frequency analysis methods, i.e. EMD and WPT combined with SVD.

3.4 Comparisons with BP and KNN

In order to verify the performance of classifier, the classifiers for K-NN and BP are compared with the multi-classifiers of SVM. The running time and the average classification accuracy of the three classifiers based on ITD-SVD are shown in Table 4. As can be seen from Table 4 that the average time consumption by SVM is 0.0106 s, and the average classification accuracy of SVM is 99.53 %, which is higher than those of the other two methods.

BP adopts the greedy strategy to search the hypothesis space, which can only obtain the local optimal solution. In the case of small samples, the performance of BP is poor. K-NN mainly solves the small sample problem by calculating the

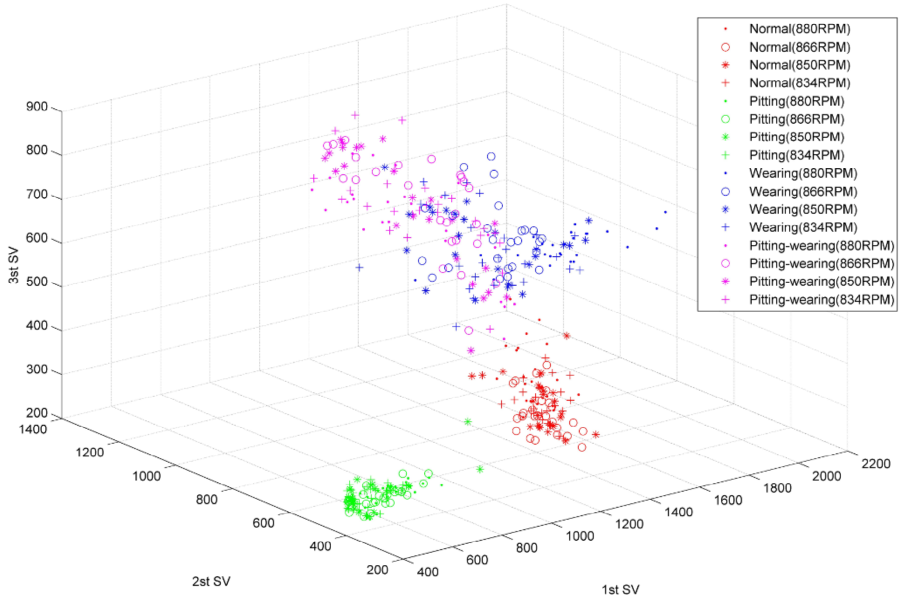

Fig. 7. Scatter plots of the three singular values extracted with EMD-SVD.

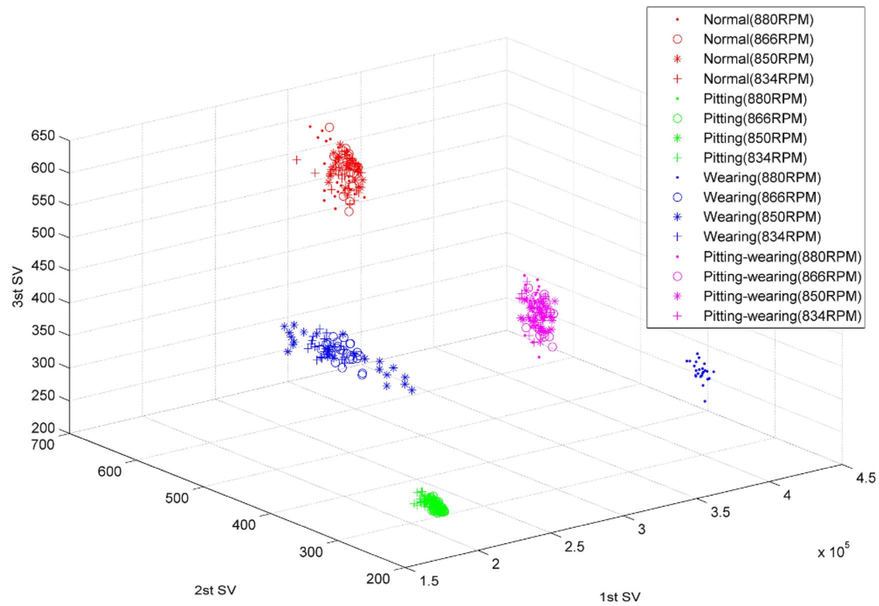

Fig. 8. Scatter plots of the three singular values extracted with WPT-SVD.

distance among samples. The computational complexity of KNN is higher than SVM. Owing to the selection of kernel function, SVM can achieve the global optimum and obtain a higher accuracy in the case of small samples. In view of this, the proposed method combining ITD-SVD with SVM can be effectively applied in gear fault diagnosis under variable conditions.

4. Conclusions

In this paper, a method based on ITD-SVD and SVM for gear fault diagnosis under variable conditions is put forward. In line with this method, the intrinsic time-scale decomposition is firstly performed to decompose the vibration signals into several PRCs, and the singular value decomposition is applied subsequently to reduce dimensionality. ITD-SVD is successfully performed to extract the fault feature vectors and improve robustness under variable conditions. A detailed comparison is made among ITD, EMD and WPT combined with SVD, and the result shows the superiority of ITD-SVD under variable conditions. Therefore, under certain circumstances, the influence of the time-varying working condition can be negligible. Finally, SVM is applied to the fault classification. With a small number of training samples, the algorithm can achieve very high accuracies. What is more, SVM is compared with K-NN and BP, as they are all regarded as ef-

Table 3. Classification results of the three feature extraction methods based on SVM.

Training samples	Testing samples	Average classification accuracy (%)		
		EMD-SVD	WPT-SVD	ITD-SVD
20	380	87.16	97.28	99.53

Table 4. Classification results of the three classifiers based on ITD-SVD.

Training samples	Testing samples	Running time (s)			Average classification accuracy (%)		
		BP	K-NN	SVM	BP	K-NN	SVM
20	380	0.2695	0.0204	0.0106	92.61	99.28	99.53

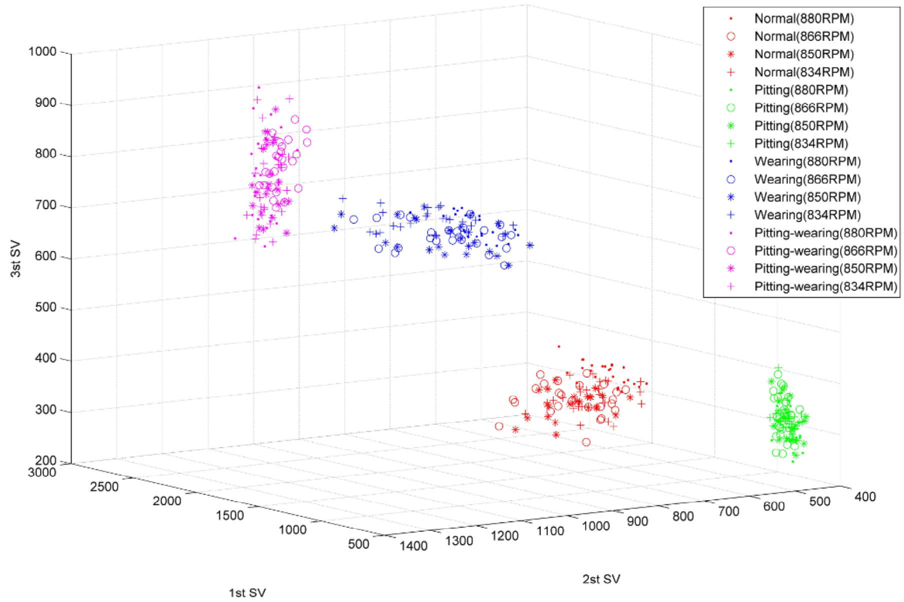

Fig. 9. Scatter plots of the three singular values extracted with ITD-SVD.

fective methods for fault classification. It is found that SVM has higher accuracy. The experimental results indicate that the proposed ITD-SVD method with SVM is suitable and efficient for gear fault diagnosis under variable conditions.

The working condition discussed in this paper is four types of speed. If the working condition is changing continuously, the accuracy of the proposed method may be influenced. In addition, the fault classification of the same fault type with different speeds is worthy of future study.

Acknowledgments

This research was supported by the National Nature Science Foundation of China (No. 61374135), the National Natural Science Foundation of Chongqing (No.cstc2016jcyjA0504) and Chongqing University Postgraduates Innovation Project (No. CYS15027). The authors are also grateful for the editors and reviewers who made constructive comments.

References

[1] L. Zhang et al., Gearbox health condition identification by neuro-fuzzy ensemble, *Journal of Mechanical Science and Technology*, 27 (3) (2013) 603-608.

[2] C. Sreepradha et al., Neural network model for condition monitoring of wear and film thickness in a gearbox, *Neural Computing and Applications*, 24 (7-8) (2014) 1943-1952.

[3] H. Liu, X. Wang and C. Lu, Rolling bearing fault diagnosis under variable conditions using Hilbert-Huang transform and singular value decomposition, *Mathematical Problems in Engineering* (2014).

[4] S. Singh and N. Kumar, Combined rotor fault diagnosis in rotating machinery using empirical mode decomposition, *Journal of Mechanical Science and Technology*, 28 (12) (2014) 4869-4876.

[5] A. Raad, J. Antoni and M. Sidahmed, Indicators of cyclostationarity: Theory and application to gear fault monitoring, *Mechanical Systems and Signal Processing*, 22 (3) (2008) 574-587.

[6] P. Boškoski and D. Juričić, Fault detection of mechanical drives under variable operating conditions based on wavelet packet Rényi entropy signatures, *Mechanical Systems and Signal Processing*, 31 (2012) 369-381.

[7] L. Hong and J. S. Dhupia, A time domain approach to diagnose gearbox fault based on measured vibration signals, *Jour-*

nal of Sound and Vibration, 333 (7) (2014) 2164-2180.
[8] C. Li et al., Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis, *Neurocomputing*, 168 (2015) 119-127.
[9] D. Zhang, D. Yu and W. Zhang, Energy operator demodulating of optimal resonance components for the compound faults diagnosis of gearboxes, *Measurement Science and Technology*, 26 (11) (2015) 115003.
[10] X. G. Tang, Y. Guo and Y. C. Zheng, Application of rolling element bearing envelope analysis based on short time Fourier transition and independent components analysis, *Journal of Mechanical Strength*, 1 (2012) 1-3.
[11] P. K. Kankar, S. C. Sharma and S. P. Harsha, Rolling element bearing fault diagnosis using wavelet transform, *Neurocomputing*, 10 (2011) 1638-1645.
[12] L. Shi, Y. Zhang and W. Mi, Application of Wigner-Ville-distribution-based spectral kurtosis algorithm to fault diagnosis of rolling bearing, *Journal of Vibration, Measurement and Diagnosis*, 1 (2011) 27-31.
[13] H. X. Zhang et al., Contrast of time-frequency analysis methods and fusion of wigner-ville distribution and wavelet transform, *Advanced Materials Research,* 805 (2013) 1962-1965.
[14] N. E. Huang, *Hilbert-Huang transform and its applications*, World Scientific (2014).
[15] L. Deng and R. Zhao, Fault feature extraction of a rotor system based on local mean decomposition and Teager energy kurtosis, *Journal of Mechanical Science and Technology*, 28 (4) (2014) 1161-1169.
[16] J. Cheng, D. Yu and Y. Yang, Application of support vector regression machines to the processing of end effects of Hilbert-Huang transform, *Mechanical Systems and Signal Processing*, 21 (3) (2007) 1197-1211.
[17] J. Zheng, J. Cheng and Y. Yang, Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis, *Mechanical Systems and Signal Processing*, 40 (1) (2013) 136-153.
[18] M. G. Frei and I. Osorio, Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences,* The Royal Society, 463 (2078) (2007) 321-342.
[19] X. An et al., Application of the intrinsic time-scale decomposition method to fault diagnosis of wind turbine bearing, *Journal of Vibration and Control*, 18 (2) (2012) 240-245.
[20] L. Bo and C. Peng, Fault diagnosis of rolling element bearing using more robust spectral kurtosis and intrinsic time-scale decomposition, *Journal of Vibration and Control* (2014) 1077546314547727.
[21] R. Jegadeeshwaran and V. Sugumaran, Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines, *Mechanical Systems and Signal Processing*, 52 (2015) 436-446.
[22] J. Tian et al., Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis, *IEEE Transactions on Industrial Electronics*, 63 (3) (2016) 1793-1803.
[23] K. Jayakumar, S. Thangavel and M. K. Elango, Back propagation algorithm for bearing fault detection of induction motor drive using wavelet pocket decomposition, *International Journal of Applied Engineering Research*, 10 (10) (2015) 26191-26208.
[24] H. Liu, X. Wang and C. Lu, Rolling bearing fault diagnosis under variable conditions using Hilbert-Huang transform and singular value decomposition, *Mathematical Problems in Engineering* (2014).
[25] J. X. Zeng et al., The de-noising algorithm based on intrinsic time-scale decomposition, *Advanced Materials Research*, 422 (2012) 347-352.
[26] Y. Tian, C. Lu and Z. L. Wang, Approach for hydraulic pump fault diagnosis based on WPT-SVD and SVM, *Applied Mechanics and Materials*, 764 (2015) 191-197.
[27] S. Suthaharan, Support vector machine, *Machine Learning Models and Algorithms for Big Data Classification*, Springer US (2016) 207-235.
[28] K. Fu et al., Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM, *Biomedical Signal Processing and Control*, 13 (2014) 15-22.
[29] D. Lee and J. Lee, Domain described support vector classifier for multi-classification problems, *Pattern Recognition*, 40 (1) (2007) 41-51.

Jianfeng Qu received the B.E. degree, M.E. degree and the Ph.D. degree in College of Automation from Chongqing University, China. Prof. Qu joined the College of Automation at the Chongqing University in 2010. He is currently an Associate Professor the College of Automation at the Chongqing University, China. His research and teaching interests include signal detection and analysis, non-stationary signal analysis, data-driven modeling and analysis.

Zhanqiang Xing received the B.E. degree in College of Automation from Chongqing University, China. He is currently a master student in the College of Automation from Chongqing University. His research interests include fault diagnosis and signal processing.