Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis

CrossMark

Ruonan Liu [a,b], Boyuan Yang [a,b], Xiaoli Zhang [a,b], Shibin Wang [a,b], Xuefeng Chen [a,b,*]

[a] The State Key Laboratory for Manufacturing Systems Engineering, Xi'an, PR China
[b] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, PR China

A B S T R A C T

Bearing plays an essential role in the performance of mechanical system and fault diagnosis of mechanical system is inseparably related to the diagnosis of the bearings. However, it is a challenge to detect weak fault from the complex and non-stationary vibration signals with a large amount of noise, especially at the early stage. To improve the anti-noise ability and detect incipient fault, a novel fault detection method based on a short-time matching method and Support Vector Machine (SVM) is proposed. In this paper, the mechanism of roller bearing is discussed and the impact time frequency dictionary is constructed targeting the multi-component characteristics and fault feature of roller bearing fault vibration signals. Then, a short-time matching method is described and the simulation results show the excellent feature extraction effects in extremely low signal-to-noise ratio (SNR). After extracting the most relevance atoms as features, SVM was trained for fault recognition. Finally, the practical bearing experiments indicate that the proposed method is more effective and efficient than the traditional methods in weak impact signal oscillatory characters extraction and incipient fault diagnosis.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Condition monitoring of rotating machinery is important for system maintenance and process automation. Roller bearing, as an important component, is not only widely used, but also a common failure unit in rotating machinery due to its complex running conditions [1]. The accuracy of mechanical system is highly dependent on the dynamic performance of bearings [2]. Severe working condition such as heavy load, high speed and lubrication oil lackage makes bearing damage inevitable [3]. Therefore, faulty rolling bearings are often the root cause of faults and they must be detected as early as possible to avoid the financial loss and disasters.

However, the measured vibration signals are always complex and non-stationary with a large amount of background noise, and the useful fault information is usually too weak to be distinguished. Furthermore, the early fault symptoms hidden in signals are always quite difficult to be found [4]. To overcome this problem, numerous vibration signal processing approaches and feature extraction methods have been developed recently including probabilistic analysis, frequency

analysis, time-domain analysis and various signal processing algorithms for the tasks of detection and diagnosis [5–7], such as Fourier transform [8], wavelet transform (WT) [9–12], empirical mode decomposition(EMD) [13–16], spectral kurtosis (SK) [17]. These methods can partition the time-frequency place and design innumerable linear time-frequency representation methods according to the requirements. However, the results of these methods are directly influenced by the base functions because signals can be analyzed very well only on the condition that the time-frequency structure of signal and base function are similar. And it is difficulty to choose proper base functions.

To avoid this problem and represent the complicate signal more efficiently, atom decomposition methods were proposed which is a signal representation method based on various parameterized basis functions library called the dictionary and have received increasing attention in signal and image processing. Matching pursuit (MP) algorithm is one of the most important atom decomposition methods which was proposed by Mallat and Zhang in 1993 [18]. At the same time, Qian and Chen proposed an idea which was similar with MP. MP is a typical greedy algorithm that chooses a waveform that is best adapted to approximate part of the signal at each iteration with a particular over-complete dictionary [19]. However, the over-complete dictionary may lead to huge computation in atom decomposition for signal processing and the decomposition results are directly influenced by the dictionary and the atoms should match with the structure characteristics of the signals [18].

Fault diagnosis often includes three steps: information acquisition, feature extraction and condition recognition. Another point of roller bearing fault diagnosis is pattern recognition [20,21]. Support vector machines (SVM) based on statistical learning theory that are of specialties for a smaller sample number and guarantee the solution are exactly same with the global optimal solution. Considering that it is hard to obtain sufficient fault samples in practice, SVM are applied to the bearing fault diagnosis because of their high accuracy and good generalization for a smaller sample number. However, it is not sensitive for incipient fault detection if use the traditional time domain indexes as inputs for SVM such as root means square (RMS), absolute average, kurtosis, square-root amplitude.
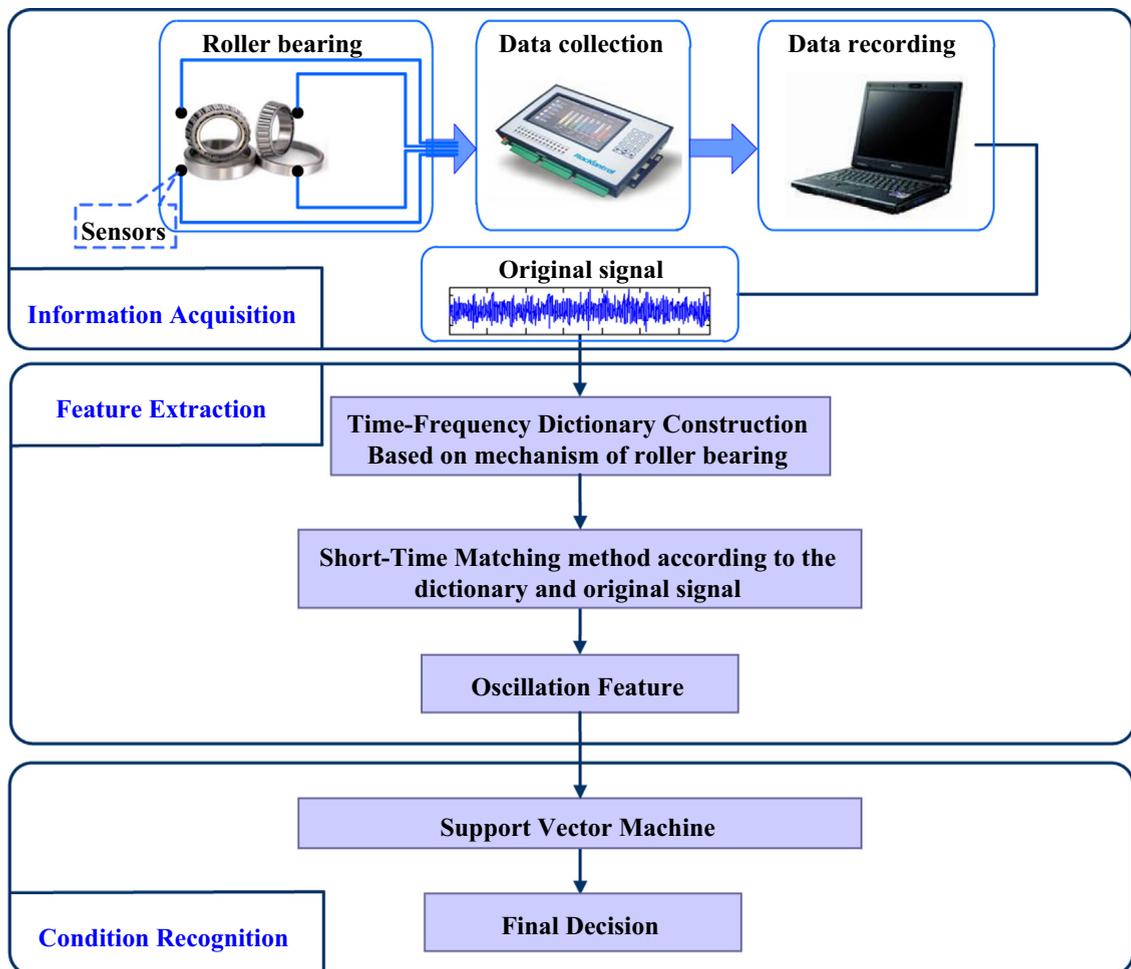

Fig. 1. Flow chart of the proposed method.

Combined with the fault characteristics of bearing and inspired by atom decomposition methods, a new fault diagnosis method based on impact time-frequency dictionary, short-time matching method and SVM is proposed to overcome the limitations of traditional methods and enhance the effectiveness of incipient fault diagnosis. First of all, according to the mechanism of roller bearing, an impact time-frequency dictionary is constructed to match the structure characteristics of signals. Then, to overcome the limitations of conventional sparse representation methods for weak impact signal oscillatory characters extraction and generate feature vectors, a short-time matching algorithm is proposed by calculating the relevancy between the time-frequency impact atoms and the signal which was intercepted by a sliding window. Finally, as a diagnostic classifier, SVM was trained by the features which were extracted by the short-time matching. Compared with the traditional signal processing and fault diagnosis methods, the proposed method can realize the detection and extraction of weak impact signal when the signal-noise-ratio (SNR) is extremely low. The process chart is shown in Fig. 1.

The remainder of this paper is organized as follows. Section 2 describes theoretical background briefly. The time-frequency atoms-driven SVM method was described in Section 3. Then, three simulations are described in Section 4 and two practical applications are proposed in Section 5. Finally, conclusions are drawn in Section 6.

2. Theoretical background

2.1. Atomic decomposition

Conventional linear signals representation theories are based on the base function decomposition. However, real signals mixed many components and cannot be represented by a base function. In recent years, many efforts and much progress have been made to improve the performance of traditional methods. And the atomic decomposition theory, which is an extension of basis function expansion, was proposed [22].

The main idea of atomic decomposition is to replace the base function sets with over-complete redundant function sets which is called the over-complete dictionary; then trace the parameterized functions matched with the signal structure features; finally signal was represented as a linear combination of few vectors which are called atoms in the dictionary. If the atoms are similar with the signal's inner structure, the signal will be accurately represented with few atoms. Therefore, in atomic decomposition theory, the sparser the decompose results are, the closer between the atoms and signals' essence and inner structure. And it is easier for feature extraction because the information is focused on that few atoms. Compared with traditional basis function expansion, atomic decomposition can represent arbitrary complex signal based on the over-complete dictionary.

Various atomic decomposition methods have been proposed to reduce the computational complexity and running time, such as frame decomposition, basis pursuit (BP) [23] and MP [18].

The problem of optimally approximating a function with a linear expansion over a redundant dictionary of waveforms is non-deterministic polynomial hard (NP-hard) [24]. Matching pursuit is an approach to represent the signal with a sub-optimal linear combination of a set of atoms selected from the dictionary. The principle of MP algorithm is described in the following. Detailed explanation can be referred to [18].

Set H as the Hilbert space. $D \in H$ is the over-complete atom dictionary and $D = \{g_\gamma(t)\}_{\gamma \in \Gamma}$. Γ is the parameters set. g_γ is the atom determined by parameter γ and $\|g_\gamma\| = 1$. Suppose that a signal f is to be analyzed and $f \in H$. With $R^0 f = f$. $R^n f$ is the residual signal after n iterations. $g_{\gamma_n} \in D$ is the best match atom in n^{th} iterations. After orthogonal projection with D, f can be decomposed as

$$f = \left\langle R^0 f, g_{\gamma_0} \right\rangle g_{\gamma_0} + R^1 f, \tag{1}$$

where $g_{\gamma_0} \in D$. $R^1 f$ is the residual signal after approximating f with g_{γ_0}.

obviously $R^1 f$ is orthogonal to g_{γ_0}, therefore

$$\|f\|^2 = \left| \left\langle f, g_{\gamma_0} \right\rangle \right|^2 + \|R^1 f\|^2, \tag{2}$$

to minimize $R^1 f$, the vector g_{γ_0} selected should satisfy

$$\left| \left\langle f, g_{\gamma_0} \right\rangle \right| = \max_{\gamma \in \Gamma_a} \left| \left\langle f, g_\gamma \right\rangle \right| \geq \alpha \sup_{\gamma \in \Gamma} \left| \left\langle f, g_\gamma \right\rangle \right|, \tag{3}$$

where Γ_0 is a set of parameters, $\Gamma_0 \in \Gamma$. α is the optimization factor and $0 < \alpha \leq 1$.

The residual $R^1 f$ is further decomposed in the same way as f. Let $R^0 f = f$, at step n, the residual signal become $R^n f$. In order to maximum $\left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n}$ should be selected by

$$\left| \left\langle R^n f, g_{\gamma_n} \right\rangle \right| \geq \alpha \sup_{\gamma \in \Gamma} \left| \left\langle R^n f, g_{\gamma_n} \right\rangle \right|, \tag{4}$$

the residual signal is then decomposed by

$$R^n f = \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^{n+1} f, \tag{5}$$

where $R^{n+1}f$ is the new residual signal after decomposed $R^n f$. In the same way, $R^{n+1}f$ is orthogonal to g_{γ_n}, therefore

$$\|R^n f\|^2 = \left| \left\langle R^n f, g_{\gamma_n} \right\rangle \right|^2 + \|R^{n+1} f\|^2 \tag{6}$$

after m iterations, the signal f can be represented as

$$f = \sum_{n=0}^{m-1} \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^m f \tag{7}$$

and

$$\|R^m f\|^2 = \left| \left\langle R^m f, g_{\gamma_m} \right\rangle \right|^2 + \|R^{m+1} f\|^2. \tag{8}$$

$$\|f\|^2 = \sum_{n=0}^{m-1} \left| \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} \right|^2 + \|R^m f\|^2 \tag{9}$$

where $R^m f$ is the residual signal after m iterations.

MP algorithm is a continuous iterative process. It will look for the atom which is most relevant to the residual signal $R^m f$ in the dictionary D.

2.2. Support vector machine

Support vector machine is a computational learning method for small samples classification based on the structural risk minimization (SRM) [25,26]. The process of SVM for a binary classification is described as follow.

Given a sample set $S = \left\{ (x_i, y_i) \right\}_{i=1}^{N}$, where x_i are the inputs and $y_i \in \{-1, +1\}$ are the labels of x_i. Each class associates with labels be $y_i = 1$ for positive class and $y_i = -1$ for negative class. Then, an optimal hyperplane $f(x) = 0$ is computed in a feature space to construct SVM and separates the given data

$$f(x) = w^T x + b = \sum_{j=1}^{M} w_j x_j + b = 0 \tag{10}$$

where $\boldsymbol{w}$ is N-dimensional vector and b is a scalar. The $\boldsymbol{w}$ and b are used to define the position of the hyperplane. $f(x)$ is the decision function to create hyperplane which will classify input data. The optimal separating hyperplane is the separating hyperplane that creates the maximum distance between the plane and nearest data, that is, the maximum margin. Fig. 2 shows an example of the optimal hyperplane.
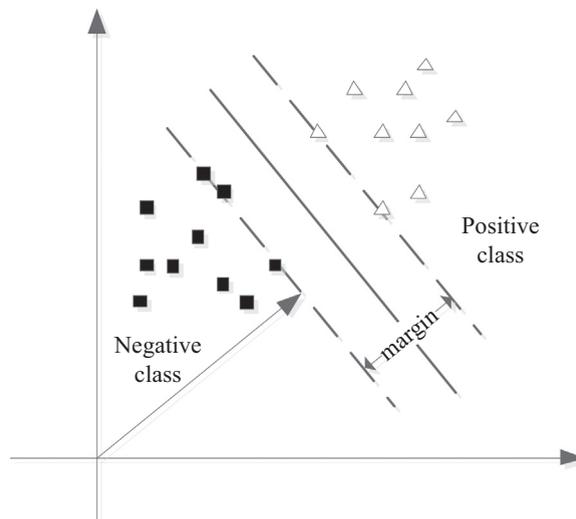

Fig. 2. The optimal hyperplane for a binary classification.

Considering the noise with slack variables ς_i and the error penalty C, the problem is to find the solution of the optimization problem:

$$\text{Minimize} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\varsigma_i, , \tag{11}$$

$$\text{Subject to}\begin{cases} y_i(w^T x_i + b) \geq 1 - \varsigma_i, & i = 1, \cdots, M \\ \varsigma_i \geq 0, & i = 1, \cdots, M \\ C \geq 0 \end{cases}, \tag{12}$$

By converting the problem with Kuhn–Tucker condition into the equivalent Lagrangian dual problem, we can get a simpler calculation:

$$\text{Minimize } L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{M}\alpha_i y_i(w \cdot x_i + b) + \sum_{i=1}^{M}\alpha_i, \tag{13}$$

We have the saddle-point equations as follows at the optimal point:

$$\frac{\partial L}{\partial w} = 0, \quad \frac{\partial L}{\partial b} = 0 \tag{14}$$

We can get:

$$w = \sum_{i=1}^{M}\alpha_i y_i x_i, \qquad \sum_{i=1}^{M}\alpha_i y_i = 0 \tag{15}$$

Using substitution Eq. (15) into Eq. (13), the dual quadratic optimization problem converts into:

$$\text{Maximize } L(\alpha) = \sum_{i=1}^{M}\alpha_i - \frac{1}{2}\sum_{i=1}^{M}\alpha_i\alpha_j y_i y_j x_i \cdot x_j \tag{16}$$

$$\text{Subject to}\begin{cases} \alpha_i \geq 0, & i = 1, \cdots, M, \\ \sum_{i=1}^{M}\alpha_i y_i = 0 \end{cases} \tag{17}$$

Using the non-linear vector function $\Phi(x_i)$, the input vectors x_i can be mapped into high-dimensional feature space. Mercer kernel returns a dot product of the feature space mappings of original data points, stated as $K(x_i, x_j) = \left(\Phi^T(x_i) \cdot \Phi(x_j)\right)$. Then, the dual quadratic optimization problem can be given by

$$\text{Maximize } L(\alpha) = \sum_{i=1}^{M}\alpha_i - \frac{1}{2}\sum_{i=1}^{M}\alpha_i\alpha_j y_i y_j K(x_i, x_j) \tag{18}$$

$$\text{Subject to}\begin{cases} \alpha_i \geq 0, & i = 1, \cdots, M, \\ \sum_{i=1}^{M}\alpha_i y_i = 0 \end{cases} \tag{19}$$

where α_i is the optimal Lagrange multipliers. Then the classifier based on the support vectors is

$$f(x) = \text{sign}\left(\sum_{i=1}^{N}\alpha_i y_i K(x, x_i) + b\right) \tag{20}$$

3. Time-frequency atoms-driven SVM method

3.1. Analysis on mechanism of roller bearing

When a roller bearing is working, the outer race usually connects with bearing pedestal; the inner roller connects with shaft and rotating with the shaft. The oscillatory system of bearing and pedestal will be stimulated to vibrate by the shaft because of structural features, assembly error and exogenous process. The mechanism of vibration is shown in Fig. 3. Therefore, in practice, the vibration signals we get from sensors are the synthesis of the vibration caused by these factors. But what we care about is only the vibration which caused by the fault. So, it is the main task that how to extrude the fault vibration from the synthesis vibration. In order to diagnosis more accurately and efficiently, the characteristics of vibration signals caused by each factor should be studied firstly.
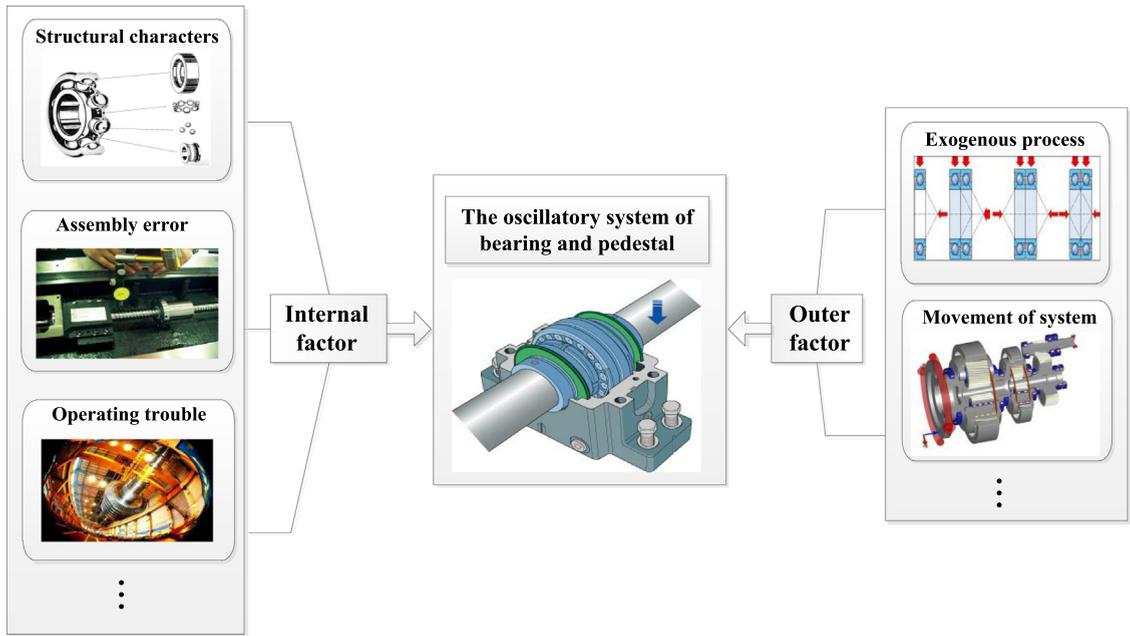
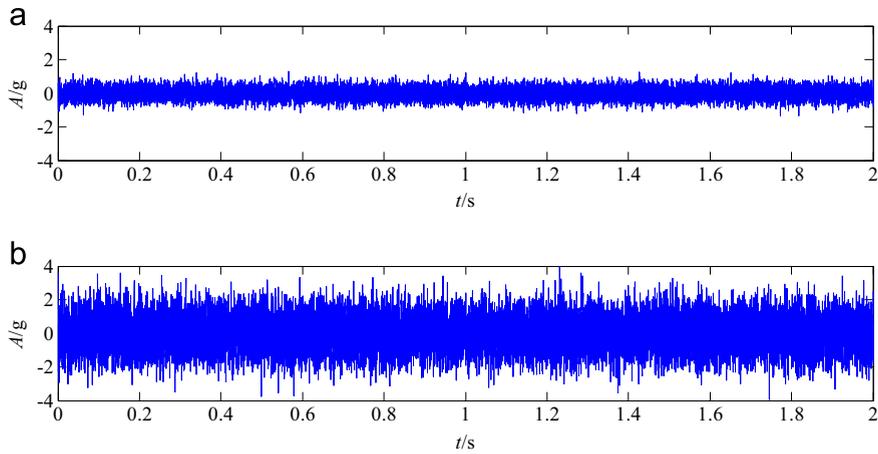Fig. 3. Mechanism of roller bearing vibration.

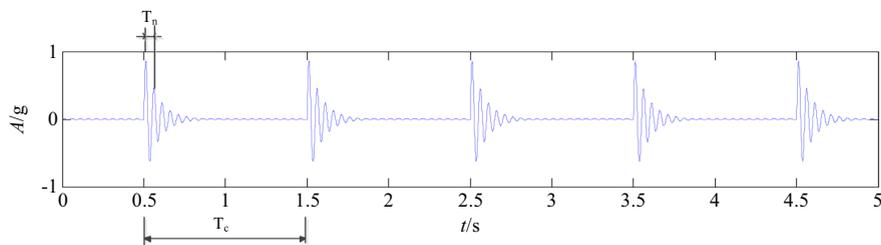Fig. 4. (a) Vibration waveform of normal bearing and (b) vibration waveform of bearing with abrasion.

Fig. 5. Vibration signals of bearing with surface damage.

The faults of bearing are mainly divided into two classes: one is abrasion, the other one is surface damage, such as pitting and peeling off.

The vibration caused by abrasion is similar with the normal bearing's vibration. All of them have strong randomness. But the vibration amplitude of abrasion is bigger than the normal ones. Fig. 4 shows the characteristics of the abrasion fault. In generally, abrasion of the roller bearing changes gradually and needs to experience a long time. It cannot lead to bearing

damage immediately, while surface damage is much more danger than it. Therefore, what we most care about is the surface damage.

If a roller bearing has a surface damage, the pulse force will suddenly change when the damage point roll from the surface of bearing. This pulse force is a wideband signal, and will cover the high frequency natural vibration frequency. Therefore, resonate will be happened, and vibration as shown in Fig. 5 will be caused.

According to Fig. 5, the components of impact vibration are divided into two parts: the first one is the vibration of low frequency. In this part, repeated impacts will be generated due to the passing of the rolling elements over the defect and the period is T_c as shown in Fig. 5. The repetition rate of the impacts, which is called as 'characteristic defect frequency', is determined by defect location, shaft speed and bearing geometry and always below 1 kHz [27]. The other one is the vibration of high frequency. This part is caused by the impact and its period is T_n as shown in Fig. 5. By means of detecting whether there exists high frequency in the vibration signal is an effective method for bearing fault diagnosis. However, the high frequency vibration of roller bearing system is very complex and has many components such as the natural vibration of sensors, rollers, inner and outer race radial bend. Therefore, how to extrude and extract the high frequency vibration signal from the comprehensive vibration for the effective operation failure criterion is the crux of the roller bearing fault diagnosis technology.

3.2. Impact time-frequency dictionary based on characteristics of roller bearing vibration signals

Decompositions of signals over family of functions which are called time-frequency atoms are well localized both in time and frequency domains and have been found many applications. The properties of decomposition results might be very different with different choices of time-frequency atoms. It is often necessary to adapt the time-frequency decomposition to the particular signal structures for feature extraction. The selection of the atom dictionary is an important point for signal sparse representation. In order to get a better performance, the over-complete dictionary should be as closely matched with signal's inner structure as possible. Then the signal will be represented with less atoms and the result will be sparser.

According to the analysis on mechanism of roller bearing and the structure characteristic of vibration signal in Section 2, this paper constructs an impact time-frequency dictionary based on the characteristics of roller bearing vibration signals. The atom g_γ is defined as:

$$g_\gamma(t) = \begin{cases} Ae^{\left(-\frac{\xi}{\sqrt{1-\xi^2}}\right) \cdot 2\pi f(t-\tau)} sin(2\pi f(t-\tau)) & t \geq \tau \\ 0 & t < \tau \end{cases} \tag{21}$$

where $\gamma = (\tau, f, \xi)$, and it is an element of the set $\Gamma = T \times F \times Z$,

$$\begin{cases} T = \{\tau_1, \tau_2, \cdots, \tau_p\} \subset R, & p \in Z^+ \\ F = \{f_1, f_2, \cdots, f_m\} \subset R^+, & m \in Z^+ \\ Z = \{\xi_1, \xi_2, \cdots, \xi_n\} \subset R^+ \cap [0, 1), & n \in Z^+ \end{cases}, \tag{22}$$

ξ is the damping coefficient, and $-\frac{\xi}{\sqrt{1-\xi^2}}$ is the damping attenuation characteristics of impact response. f is the frequency of the atom. ξ and f vary from different atoms. τ is the initial moments of impact response. A is used for atom normalization. Let H refer to the Hilbert space, $D \subset H$ is the over-complete dictionary for sparse representation. Suppose $g(t)$ is real and continuously differentiable. Therefore, the family $D = \left(g_\gamma(t)\right)_{\gamma \in \Gamma}$ is extremely redundant. Fig. 6 illustrated the waveform of the atoms. According to Fig. 6, we can see that the waveforms vary from different ξ, f, τ. The waveform will oscillates faster as ξ increase with other parameter unaltered; the waveform will decrease faster as f increase with other parameter unaltered; and the waveform will be translated as τ changes.

3.3. Oscillatory feature extraction based on sliding window and the correlation coefficient

In MP algorithm, relevance of the residual signal and the atom is measured by the inner product. However, the inner product will be influenced by the l^2-norm of the signal. What's more, signal was represented as a linear combination of few atoms. But what we care about is the extraction of impact signal and representation of a linear combination of atoms can lead to huge computational cost.

Therefore, in this paper we proposed a novel feature extract method based on impact time-frequency dictionary, correlation filter [28] and short-time window concept.

In this method, we define the correlation coefficient κ_γ as:

$$\kappa_\gamma = \cos\theta = \frac{\left|\langle g_\gamma, x(t)\rangle\right|}{\|g_\gamma\|_2 \|x(t)\|_2} \tag{23}$$

Considering $\gamma \in \Gamma, \kappa_\gamma$ should be a multidimensional matrix and its dimension is decided by the space $\Gamma = T \times F \times Z$. In this way, the influence of the l^2-norm of the signal can be eliminated. What's more, we use a short-time signal $f(\tau, \tau + W_s)$ instead of the entire signal $x(t)$ to match the dictionary D, where W_s is the atom support width, and τ is the initial moments

of the short-time signal. With smaller W_s, the correlation coefficient κ_γ can be bigger. However, the value of W_s should at least no less than the width of impact oscillation waveforms. Therefore, in the practical applications, the width of impact oscillation waveform should be estimated by observing and studying the time domain waveforms of signals. And the value of W_s can be determined according to the width of impact oscillation waveform. Then, the parameters of D become f, ξ instead of τ, f, ξ. And in order to find the atom g_γ which is most relevant to $x(t)$ at moment τ, the maximum of correlation coefficient matrix κ_γ should be found:

$$\kappa(\tau) = \max_{\substack{f \in F \\ \xi \in Z}} \kappa_\gamma^\tau = \kappa\left\{\tau, \overline{f}, \overline{\xi}\right\}, \tag{24}$$

where κ_γ^τ is the subset of κ_γ at moment τ, $\overline{f}, \overline{\xi}$ is the frequency and damping coefficient of $\kappa(\tau)$. The calculation of $\kappa(\tau)$ is to seek for the peak of the space curved surface at moment τ, as shown in Fig. 7

$$P_\tau = \left\{\kappa_\gamma^\tau(f, \xi) : f \in F, \xi \in Z\right\}. \tag{25}$$

After sliding the entire signal $x(t)$, we can get a curve of $(\tau, \kappa(\tau))$. This curve has multiple peaks and unilateral oscillatory damping behavior with the frequency and damping coefficient $\overline{f}, \overline{\xi}$. Suppose $(\tau', \kappa(\tau'))$ is the maximum point of $\kappa(\tau)$. This
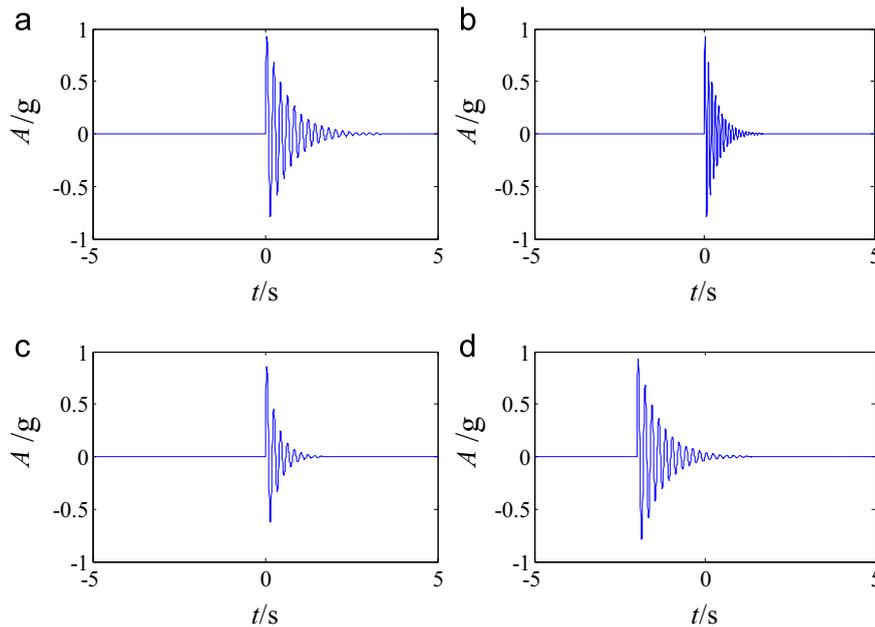

Fig. 6. Waveform of the atoms (a) g_γ with $\xi = 0.05, f = 5, \tau = 0$; (b) g_γ with $\xi = 0.1, f = 5, \tau = 0$; (c) g_γ with $\xi = 0.05, f = 10, \tau = 0$; (d) g_γ with $\xi = 0.05, f = 5, \tau = 2$.
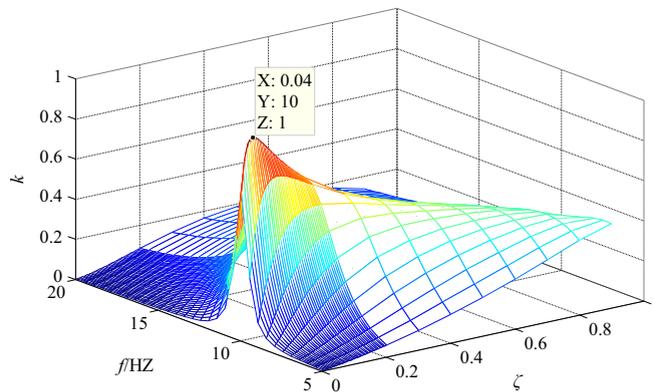

Fig. 7. The space curved surface of correlation coefficient κ_γ^τ of signal $f(t)$ and atom $g_\gamma(\tau = \tau')$.

point indicates there is a section of signal beginning from the moment τ' matches best with corresponding atom g_γ. Damping coefficient ξ and frequency f of the corresponding atom in $(\tau', \kappa(\tau'))$ represent the oscillation characteristics of the impact.

3.4. Fault recognition based on SVM and time-frequency impact atoms

We can extract n atoms with the n biggest correlation coefficient κ_γ when a signal was composed of n types of impact signal. n can be determined by the number of impact components or heuristic method. When κ_γ is smaller than a certain extent, the atoms are matching with more noise than signal. Therefore, in a certain range, the larger n is, the more atoms are used for diagnosis and the result will be better. However, the run time will increase as well if n gets larger. Therefore, n should be determined properly by balancing the run time and diagnostic accuracy. These n atoms represent the characteristics of impact's properties in the operation of the roller bearing.

After extracting the n most relevant atoms, it was found that the damping coefficient extracted by the proposed method is not as precise as the frequency coefficient due to the influence of noise when SNR is low. Therefore, parameter f is more efficient and suitable. Then, f of these atoms can be used as the input features of SVM, and faults can be detected in the incipient stage.

4. Simulated study

To verify the validity of this method in the present study, three simulation experiments were conducted. It has been found that the proposed method is more efficient, and can extract the impact signal when the signal-noise-ratio (SNR) is extremely low. The frequencies of the impact components can also be acquired by the parameters of atoms. Firstly, a simulation was performed to verify the tolerance to noise of the short-time matching method. Secondly, another simulation was performed to illustrate the effectiveness of this method in the weak signal detection and impact signal extraction with a comparison of MP. Thirdly, a gradual increasing impact signal with invariable harmonic and Gaussian noise amplitude is simulated to verify the performance of the method. In these simulations, $s(t)$ and $h(t)$ represent a single impact signal and a single harmonic signal respectively, which are defined as:

$$s_j(t) = \begin{cases} 0 & t < 0 \\ \exp\left(-\dfrac{\xi_j}{\sqrt{1-2\xi_j^2}} \cdot 2\pi f_j t\right) \cdot \sin\left(2\pi f_j t\right) & t \geq 0 \end{cases} \tag{26}$$

$$h(t) = \begin{cases} 0 & t < 0 \\ \sin\left(2\pi f_h t\right) & 0 \leq t < 0.05 \,, \\ 0 & t \geq 0.05 \end{cases} \tag{27}$$

4.1. Simulation 1

In this subsection, the proposed method is applied to extract weak impact signal from the original signal polluted by strong noise. Two periodic impact components are used to simulate the fault signal generated by two faults. The simulation signal is defined as:

$$x(t) = x_1(t) + x_2(t) + x_3(t), \tag{28}$$

$$\begin{cases} x_1(t) = s_1\left(t - T_0 - [t/T] \times T\right) \\ x_2(t) = s_2\left(t - T'_0 - [t/T'] \times T'\right), \\ x_3(t) = n(t) \end{cases} \tag{29}$$

where $s_1(t)$ and $s_2(t)$ are defined in Eq. (26) as:

$$s_1(t) = \begin{cases} 0 & t < 0 \\ \exp\left(-\dfrac{\xi_1}{\sqrt{1-2\xi_1^2}} \cdot 2\pi f_1 t\right) \cdot \sin(2\pi f_1 t) & t \geq 0 \end{cases} \tag{30}$$

$$s_2(t) = \begin{cases} 0 & t < 0 \\ \exp\left(-\dfrac{\xi_2}{\sqrt{1-2\xi_2^2}} \cdot 2\pi f_2 t\right) \cdot \sin(2\pi f_2 t) & t \geq 0 \end{cases} \tag{31}$$

with $T_0 = 0.025s, T'_0 = 0.125s, T = 0.19s, T' = 0.2s, \xi_1 = 0.07$, $f_1 = 200$ Hz, $\xi_2 = 0.07$, $f_2 = 333$ Hz. Fig. 8 shows the waveform of simulation signal. $x_1(t)$ and $x_2(t)$ are two periodic impact signals as shown in Fig. 8(a) and (b), $x_3(t)$ is Gaussian white

noise with zero-mean and 1 standard deviation, SNR=0.1023 dB, as shown in Fig. 8(c). Fig. 8(d) displays the waveform of mixed signal.

Fig. 9 displays the detection result of the simulation signal by using the proposed method. Fig. 9(a) shows the correlation coefficient κ_γ of mixed signal and the dictionary D. Fig. 9(b) shows the reconstructed signal using the five biggest κ_γ and corresponding atoms. The result reveals that the proposed method can eliminate the influence of irrelative noise and effectively extract the target signal submerged in strong noise.
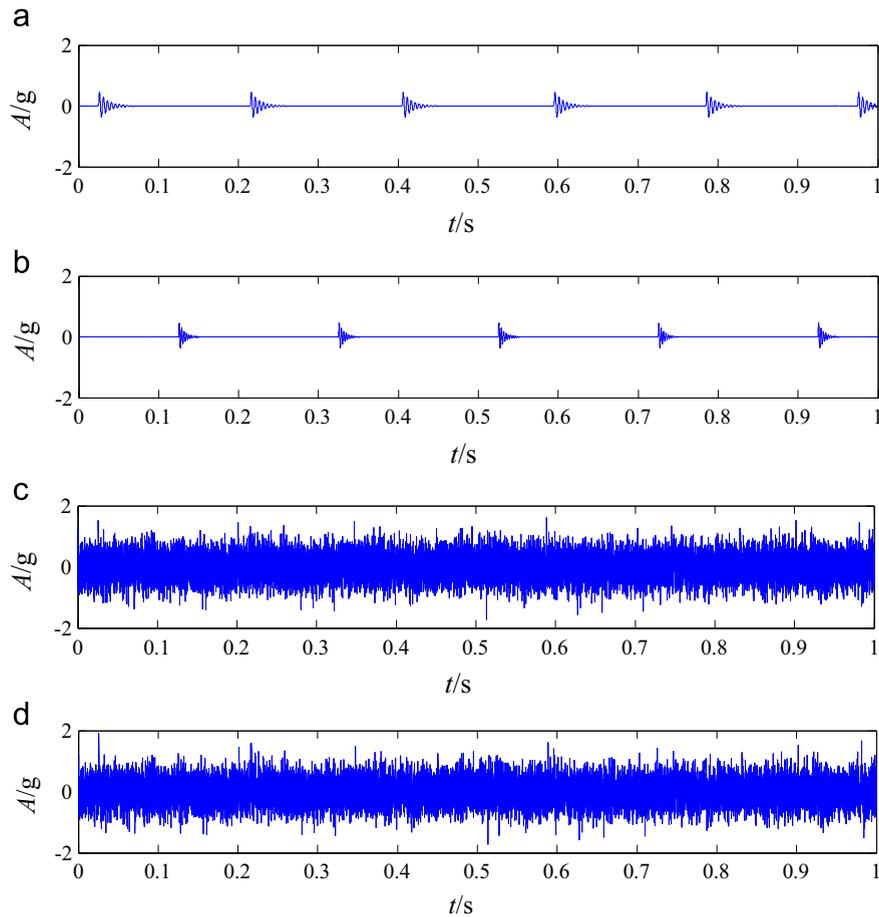
Fig. 8. The constructed signal: (a) the waveform of periodic impact component $s_1(t)$ with $\xi_1 = 0.07$, $f_1 = 200$ Hz; (b) the waveform of periodic impact component $s_2(t)$ with $\xi_2 = 0.07$, $f_2 = 333$ Hz; (c) the noise signal $n(t)$; (d) the waveform of mixed signal.
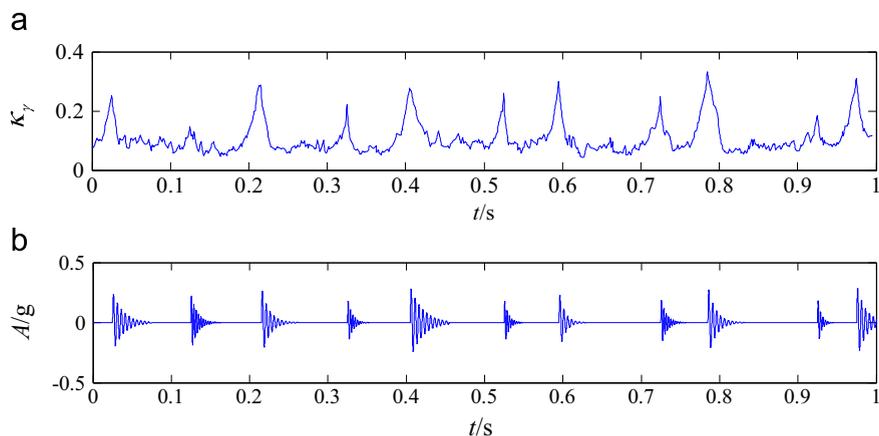
Fig. 9. The simulation results: (a) the correlation coefficient κ_γ; (b) the reconstruction impact signal using the proposed method.
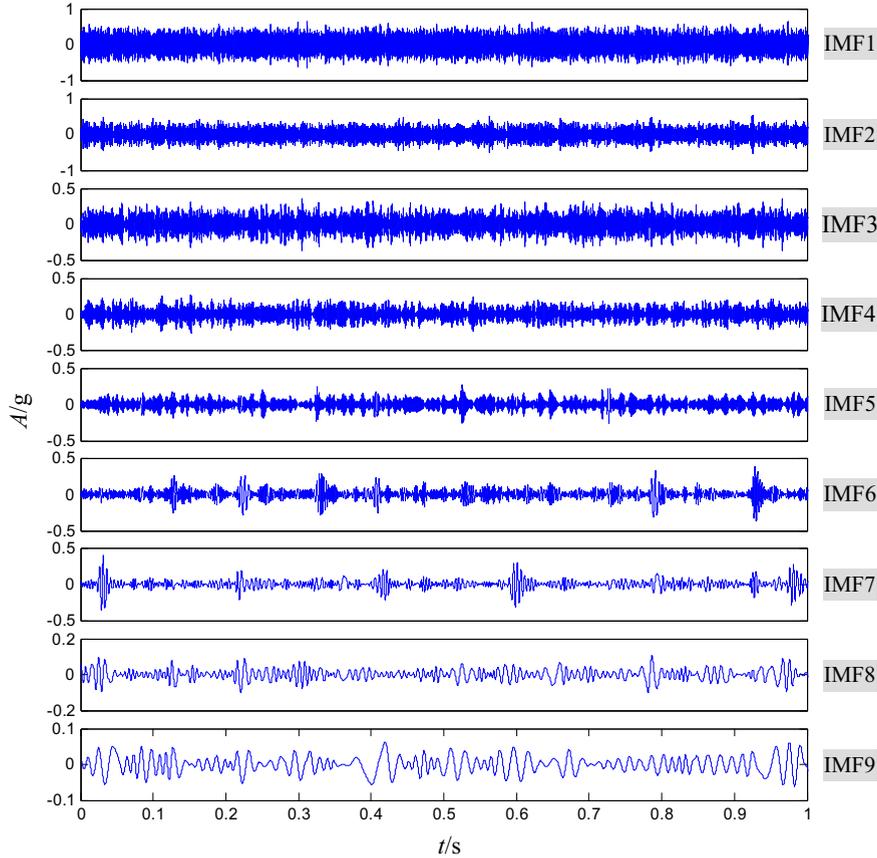
Fig. 10. The decomposition result of the simulation signal by using EMD.

For purpose of comparison, the same simulation signal is analyzed by EMD and spectral kurtosis (SK). Fig. 10 illustrates the detection result by using the EMD. And it can be found that there is no explicit periodic transients can be extracted because of the influence of strong noise. Additionally, the kurtogram of SK and the resulting filter signal are illustrated in Fig. 11. It can be found that the results are not obvious enough and fail to extract the periodic impact signal completely.

4.2. Simulation 2

In this simulation, the proposed method is applied to extract impact signal from the original signal with sine signal and noise. The simulated signal is defined as:

$$x(t) = x_1(t) + x_2(t) + x_3(t) \tag{32}$$

$$\begin{cases} x_1(t) = s\left(t - T_0 - [t/T] \times T\right) \\ x_2(t) = h\left(t - T_0' - [t/T'] \times T'\right), \\ x_3(t) = n(t) \end{cases} \tag{33}$$

with $T_0 = 0.025s, T_0' = 0.125s, T = T' = 0.2s, \xi_1 = 0.07, f_1 = f_h = 200$ Hz. $s(t)$ and $h(t)$ are described in Eqs. (26) and (27). x_1 is a periodic impact signal as shown in Fig. 12(a), x_2 is a cyclic occurring harmonic signal as shown in Fig. 12(b), x_3 is Gaussian white noise with zero-mean and 1 standard deviation, SNR=11.3150 dB, as shown in Fig. 12(c). Fig. 12(d) shows the simulation signal x which is constructed by Eq. (32).

The simulated signal is processed to match the time-frequency impact dictionary by the proposed method. The dictionary is $D = \left(g_\gamma(t)\right)_{\gamma \in \Gamma}$ which is constructed in 3.2 with damping coefficient ξ, frequency of the atom f_0, where $\xi \in [0.05, 0.15]$, $f \in [100, 500]$. Fig. 13 displays the detection result of the simulated signal. Fig. 13(a) is the impact component in original impact signal. Fig. 13(b) is the correlation coefficient κ_γ and Fig. 13(c) is κ_γ after smoothing. Fig. 13(d) is the reconstructed signal using the five biggest κ_γ and corresponding atoms. It is evident that the impact signal is completely and purely extracted from the mixed signal. By reconstructing signal, the impact vibration signal caused by fault can be extracted successfully. What's more, frequency of the impact signal can also be obtained from the corresponding atoms.
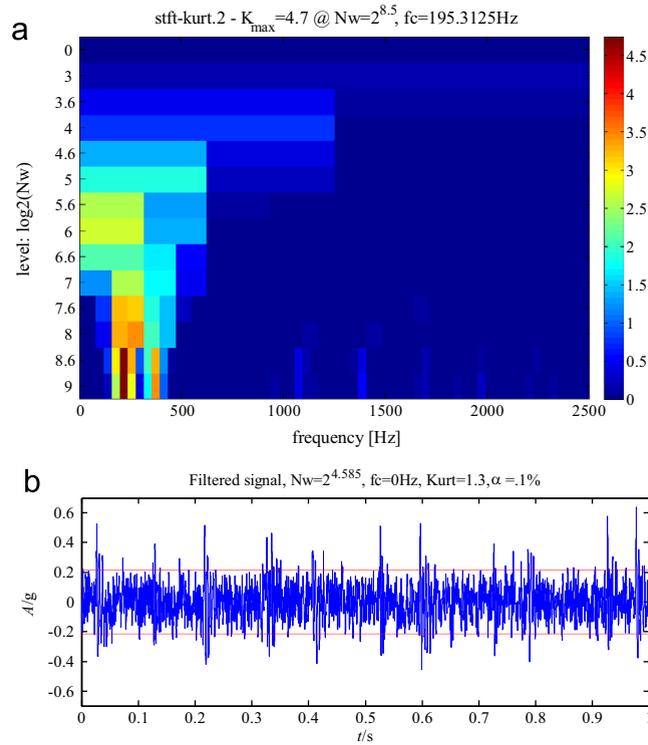
Fig. 11. The analysis result of the simulation signal by using SK: (a) the kurtogram; (b) the resulting filter signal.
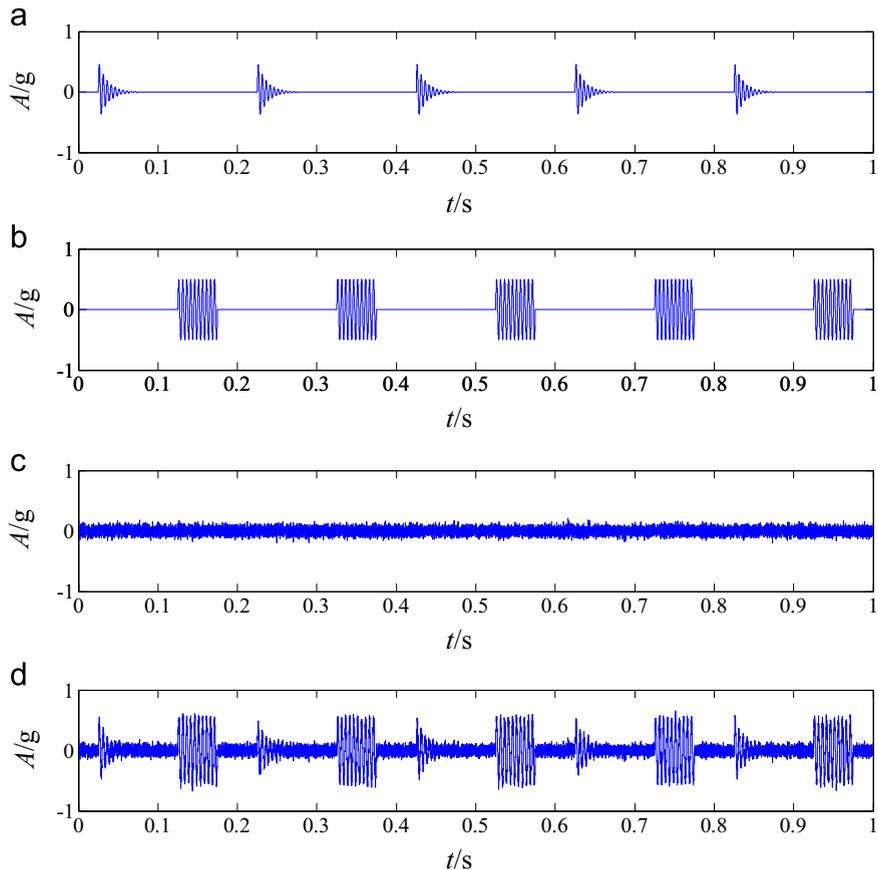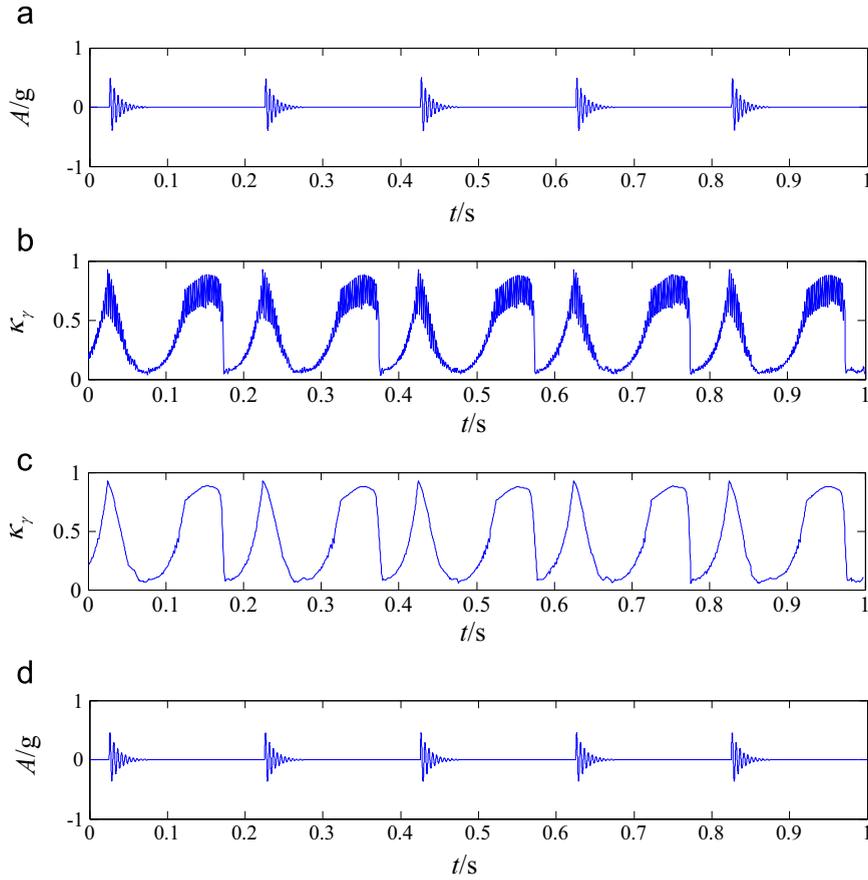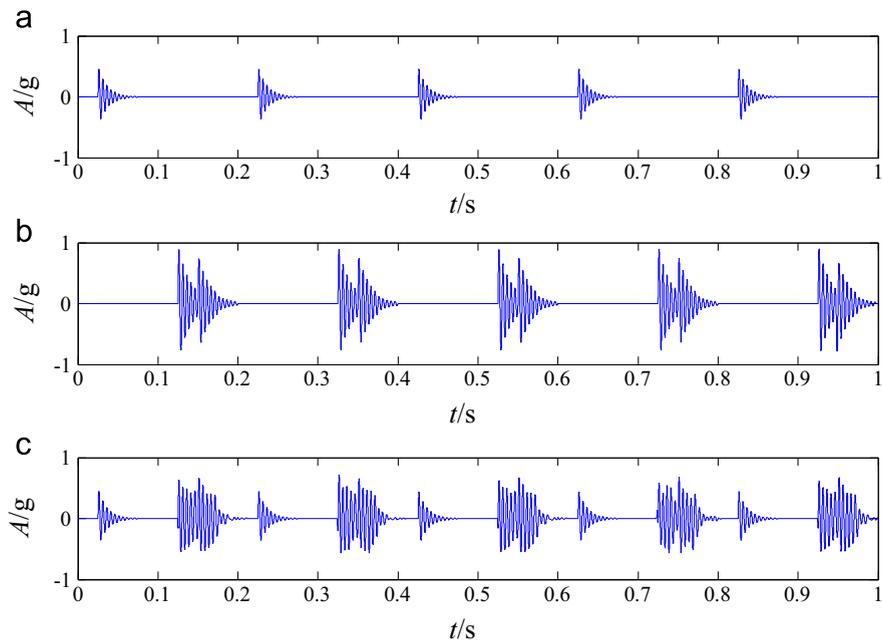

Fig. 12. Waveforms of three simulated components and the simulation signal: (a) the periodic impact signal x_1; (b) the periodic sine signal x_2; (c) Gaussian white noise x_3; (d) the simulated signal x.

Fig. 13. The detection result of the proposed method: (a) the original impact signal; (b) the correlation coefficient κ_γ; (c) κ_γ after smoothing; (d) the reconstruct signal.

Fig. 14. The detection result of MP method: (a) the original impact signal; (b) the result after 10 iterations; (c) the result after 30 iterations.
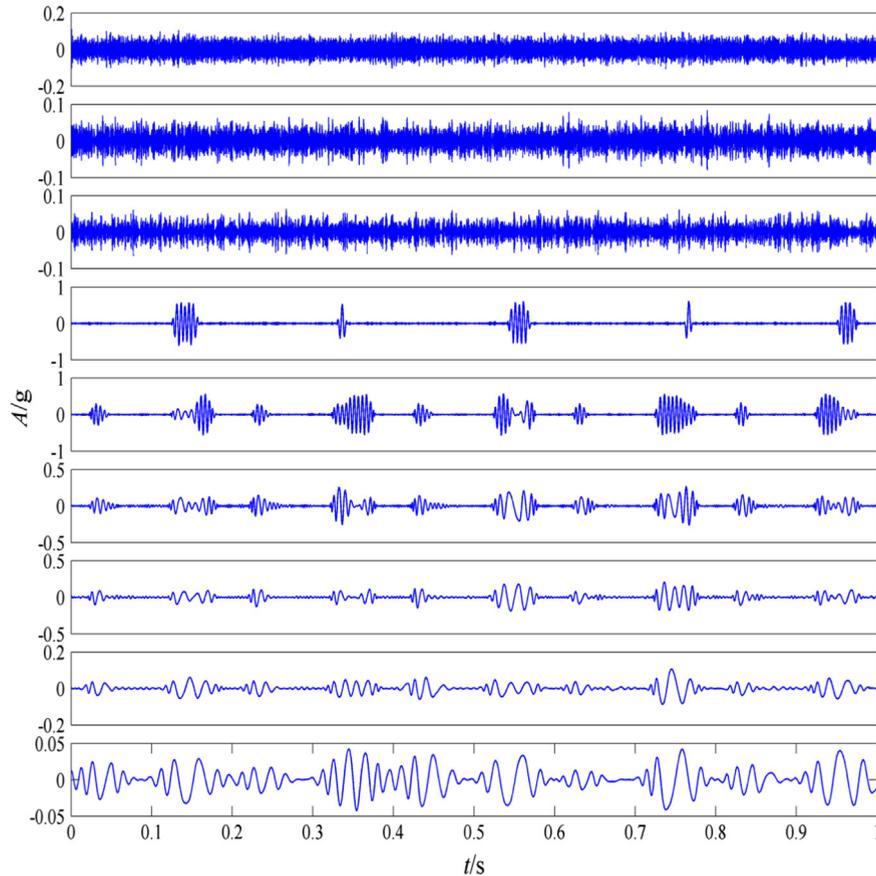
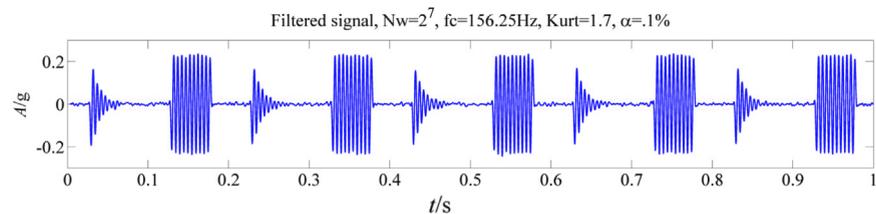**Fig. 15.** The decomposition result of the simulation signal by using EMD.

Fig. 16. The analysis result of the simulation signal by using SK.

Fig. 14 illustrates the detection result of the same simulation signal by MP method and the same dictionary D. Fig. 14(a) shows the impact component in original signal. Fig. 14(b) shows the reconstruct signal with 10 iterations. Fig. 14(c) shows the reconstruct signal with 30 iterations. It can be found that MP method matches harmonic signal first and the impact signal was matched after 30 iterations even though atoms in the dictionary D are all impact signals, because the correlation was measured by the inner product, while a maximum inner product does not lead to the most relevant impact component. The same simulated signal is also analyzed by EMD and kurtogram. The analysis result of EMD is shown in Fig. 15. It can be found that there is no explicit impact component being extracted. And the resulting filter result of kurtogram is shown in Fig. 16. It can be found that kurtogram can extract impact signal and harmonic signal from the noisy signal. However, it cannot extract only impact signal from the original signal. The comparison results illustrate the excellent performance of the proposed method in the application of extracting impact signal from the original interfered signal.

What's more, the efficiency is also improved. Table 1 shows the run times of two methods. From Table 1 we can see that the run time of MP is 18 times more than the short-time matching method.

The above simulation results demonstrate the effectiveness and high efficiency of the proposed method in the extraction of specific signal, in this study is the impact component, while MP method is for purpose of signal sparse representation. MP is a greedy algorithm which is to find the biggest inner product. However, the most relevant atom not always leads to

Table 1
Run time of the two methods.

Method	Run time (s)
Short-time matching method	0.9207
MP with 10 iterations	5.7930
MP with 30 iterations	16.6232

biggest inner product. While in the proposed method which is based on the correlation coefficient κ_γ, the impact signal was matched firstly.

4.3. Simulation 3

In this part, a simulated signal is used to verify the effectiveness of the proposed method. In actual vibration signal, the harmonic signal is always existed while the impact part will appear and increase after operating for some time. What's more, the aperiodic impulse noise can also appear in most cases. Therefore, we define the constructed signal as:

$$x(t) = x_1(t) + 0.05x_2(t) + 0.3x_3(t) + x_4(t) \tag{34}$$

where

$$x_1(t) = \begin{cases} 0 & i \leq 20 \\ 0.05 * (i-20) * s_1(t - T_0 - [t/T] \times T) & 20 < i \leq 40 \end{cases} \tag{35}$$

$$x_2(t) = \sin(2\pi \times f_h \times t), \tag{36}$$

$$x_3(t) = n(t), \tag{37}$$

$$x_4(t) = \begin{cases} 0 & 0 \leq t < \tau \\ s_2(t) & \tau \leq t < \tau + 0.05 \\ 0 & t \geq \tau + 0.05 \end{cases} \tag{38}$$

with $T_0 = 0.025$s, $T = 0.05$s,$\xi_1 = 0.05$, $f_1 = 1000$ Hz, $f_h = 200$ Hz, τ is a random value in $[0, 0.95)$. i is sample number and $0 \leq i \leq 40$. We assume that the first 20 samples are normal and without impact part and the last 20 samples with impact signal. x_1 is a periodic increased impact signal, which is increasing from 0 to 1 at 0.05 intervals. x_2 is a continuous harmonic signal, x_3 is Gaussian white noise with zero-mean and 1 standard deviation. When $i = 40$, the waveforms of them are shown in Fig. 17. x_4 is the aperiodic impulse noise with $\xi_2 = 0.02$ and $f_2 = 3000$ Hz, which is added randomly in each sample and lasts 0.05 s. τ is the start time of aperiodic impulse noise. $s_1(t)$ and $s_2(t)$ are described in Eq. (26). The waveform of $s_2(t)$ is shown in Fig. 18.

Root means square (RMS) of the signal are shown in Fig. 19.

Then x is processed to match the dictionary D. The parameters are: damping coefficient $\xi = \{0.005 : 0.01 : 0.2\} \cup \{0.2 : 0.1 : 0.9\}$, frequency of the atom $f \in [100 : 10 : 1500]$. The range of damping coefficient ξ has almost covered the entire range and is divided non-equally to ensure the high resolution when ξ is low. Many practical applications show that it is a interference fit or a tight transition fit between inner race and axis, while it is a loose transition fit between outer race and bearing pedestal. And sensors are usually placed in bearing pedestal to measure the vibration. As a result, when the failure of bearing emerges, no matter it is in inner race, outer race or bearing rollers, the inherent vibration of outer race will be risen and also generate modulation phenomenon. Therefore, frequency coefficient f represents the outer race first-order inherent frequency and the range of f can be determined by the spectrogram of corresponding bearing. Figs. 20 and 21 display the waveforms of 4 most relevant atoms for different samples. Fig. 20 is the 8th sample. Fig. 21 is the 27th sample.

From Fig. 20, we can see that in 8th sample, when there is no impact component, the main matching atoms are harmonic or the impact which is similar to harmonic with small f. In 27th sample, all main matching atoms are obvious impacts with large f.

Then we use f as the inputs of SVM classifier. The predict results based on SVM and different features are shown in Fig. 22.

From the classifier results in Fig. 22, we can see that the impact component was detected in 26th sample, when the impact amplitude is 0.3, which energy is 6.72% of the total energy.

For purpose of comparison, the 26th simulated signal, when the incipient weak impact component is identified by the proposed method, is analyzed by EMD and SK respectively. Fig. 23 shows the IMFs derived from EMD. We can see that there is no explicit incipient impact component being observed. The resulting filter signal of SK is shown Fig. 24. It can be observed that the SK is also insensitive to weak impact signal. The simulation results demonstrate the effectiveness of the proposed method in impact recognition and condition classifier.
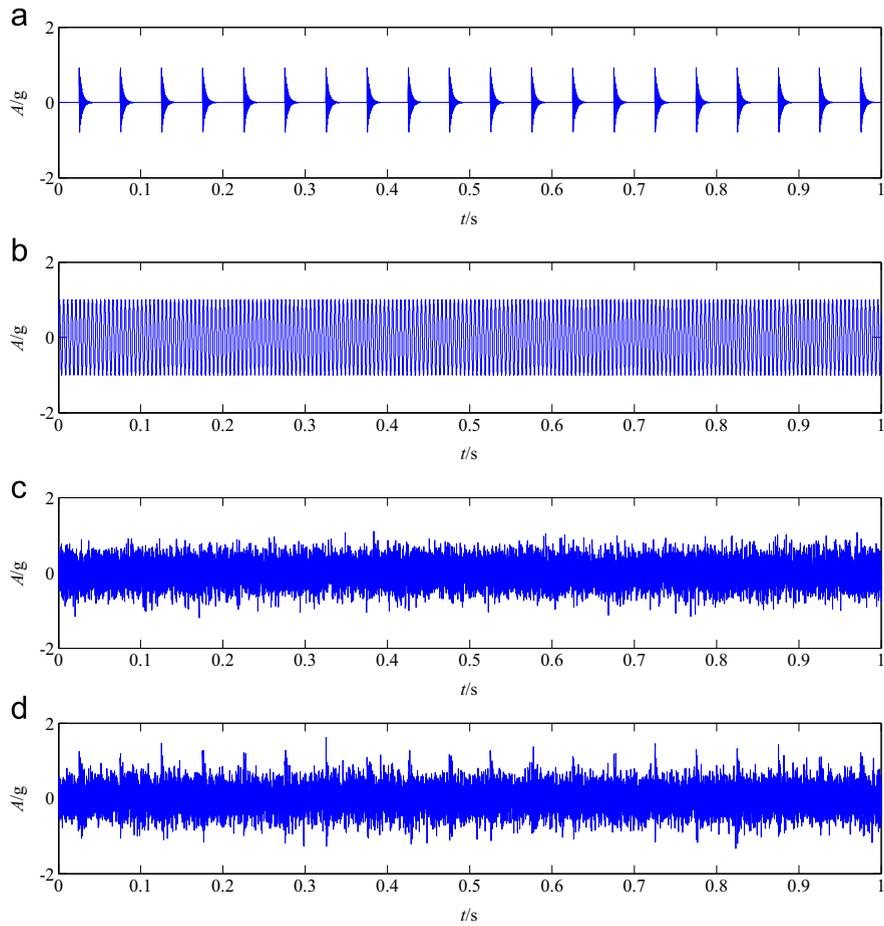
Fig. 17. The waveforms of: (a) the periodic increased impact signal x_1; (b) the continuous harmonic signal x_2; (c) the white noise x_3; (d) the constructed signal x.
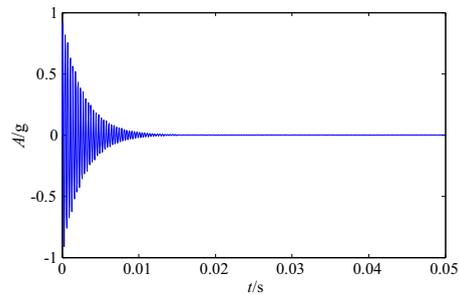
Fig. 18. The waveform of the aperiodic impulse noise x_4.
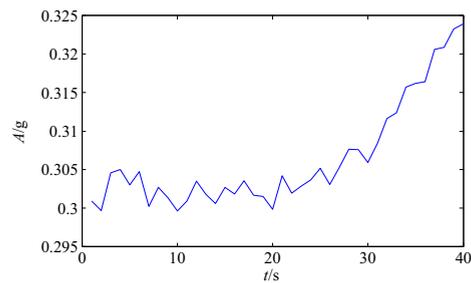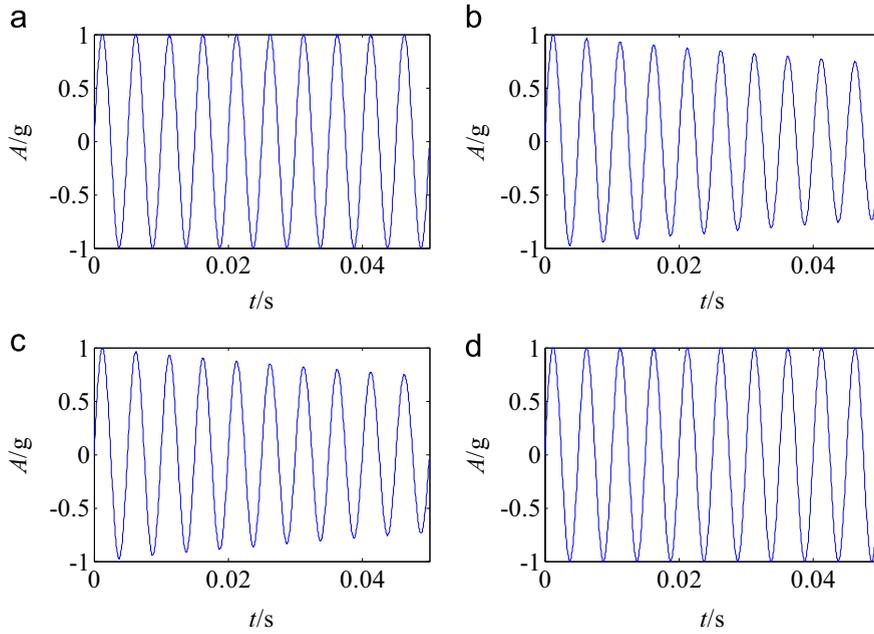

Fig. 19. The waveform of RMS.

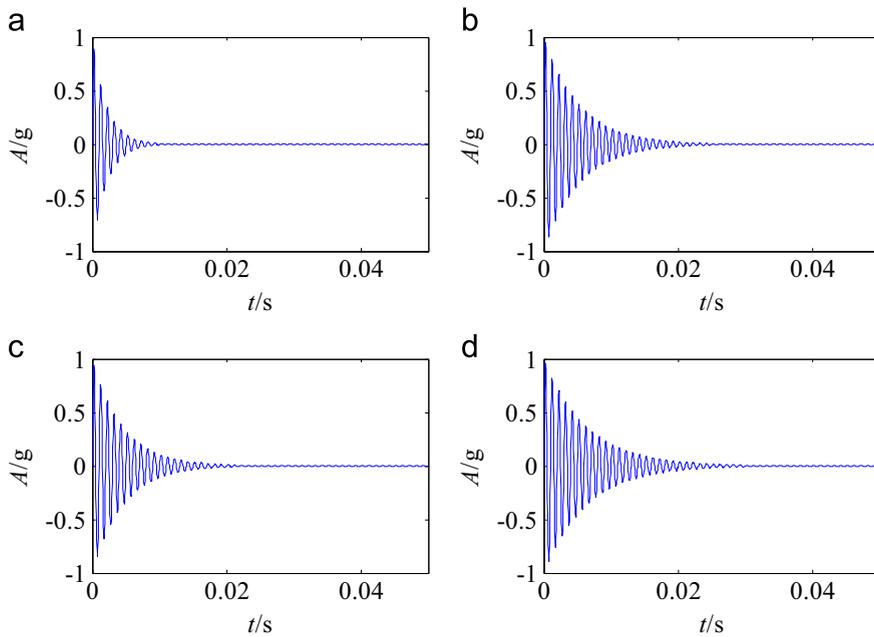**Fig. 20.** Waveforms of the 4 most relevant atoms for the 8th sample.

Fig. 21. Waveforms of 4 most relevant atoms for the 27th sample.

5. Applications to practical bearing vibration signals

To show the effectiveness of the proposed method for bearings incipient fault diagnosis, the applications to two run-to-failure bearing life test are studied in this section.

5.1. Case 1

In this subsection, a roller bearing life test was used to prove the validity of the theoretical analysis and the ability for incipient fault detection. To further demonstrate the effectiveness of the proposed method, the results of fault feature extraction by the proposed method are compared with traditional time-domain features.
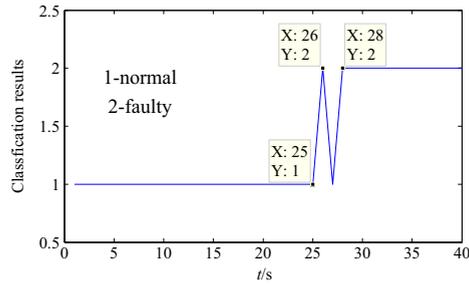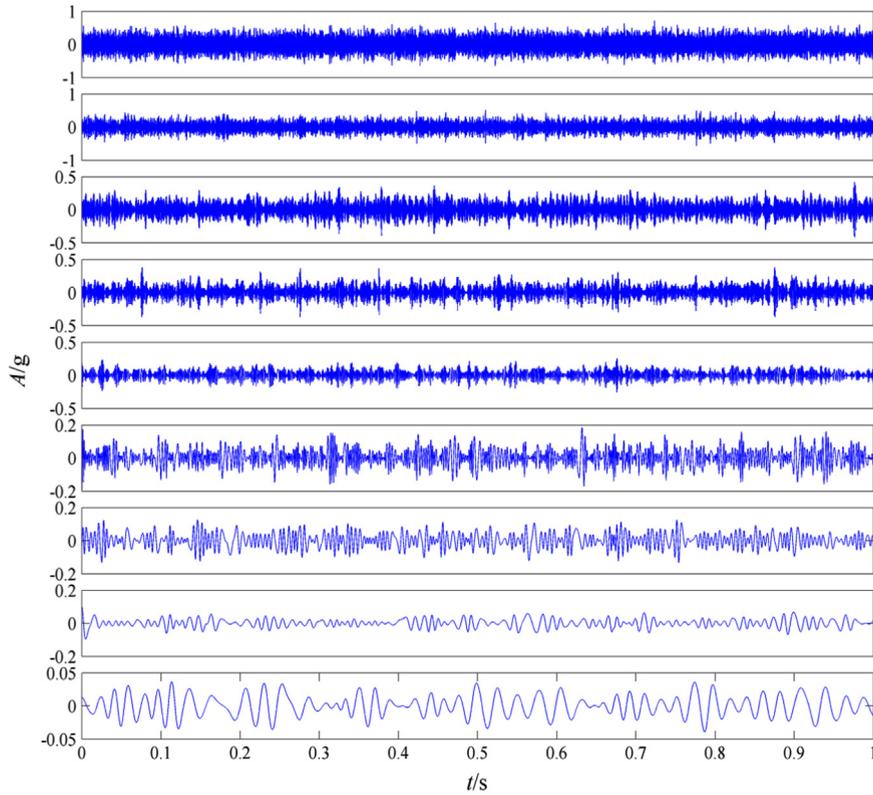
Fig. 22. Classifier results.

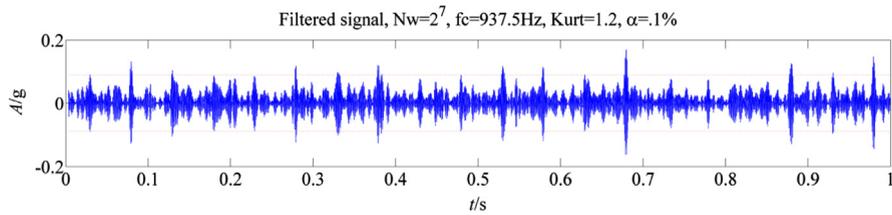**Fig. 23.** The decomposition result of the simulation signal by using EMD.

Fig. 24. The analysis result of the simulation signal by using SK.

5.1.1. Experiment introduction

This bearing run-to-failure test data are shared by the Center of Intelligent Maintenance System (IMS) in the University of Cincinnati [29].The test was performed under normal load conditions on a specially designed test rig which is shown in Fig. 25.

The bearing test rig hosts four test bearings on one shaft which is driven by an AC motor and coupled by rub belts. The rotation speed of the bearing was kept constant at 2000 rpm. A radial load was added to be 600 lbs. The vibration signals are collected every 10 min with a sampling frequency of 20 kHz per channel by a National Instruments LabVIEW program. The length of data in each sampling is 20,480 points. In the entire life span, a total of 984 groups of data were collected. At the end of the test-to-fail experiment an outer race failure occurred on bearing 1.

For the purpose of comparison, four time-domain features of the same signal were also used as inputs of SVM. They are root means square (RMS), absolute average, kurtosis, square-root amplitude. The waveforms of these features are shown in Fig. 26.

5.1.2. Feature selection

After information acquisition, short-time matching algorithm was utilized for feature extraction.

The dictionary we used is $D = \left(g_\gamma(t)\right)_{\gamma \in \Gamma}$ which is constructed in 3.2 with damping coefficient ξ, frequency of the atom f_0 and the atom support width W_s. According to the intrinsic frequency of outer race and the width of shock response waveform, the parameter space Γ is determined as $\xi = \{0.005 : 0.01 : 0.2\} \cup \{0.2 : 0.1 : 0.9\}$, $f = \{500 : 5 : 5000\}$, $W_s = 0.004$.

Then, the vibration signal is processed by the proposed method. Fig. 27(a) is an original signal in 100 h. Fig. 27(b) is the correlation coefficient κ_γ. In this experiment, we extract the first 100 biggest κ_γ and the corresponding atoms. Fig. 28 displays the waveforms of 9 atoms with the first 9 biggest correlation coefficients κ_γ in 18.33 h, when the roller bearing was normal. The waveforms of 9 atoms with the first 9 biggest correlation coefficients κ_γ in 88.83 h, when the roller bearing began to fault but is not obvious, are shown in Fig. 29. And the waveforms of atoms in 116.67 h, when the fault is obvious, are shown in Fig. 30.

From the waveforms of atoms at different moments and bearing conditions, we can see that the atoms can reflect the bearing conditions accurately. When the bearing is normal, the waveforms of the most relevant atoms are similar to harmonic with smaller damping coefficient ξ and frequency f because the original signal is mainly constituted by harmonic components. In 88.83 h, when in fault initial phase, original signal began to add impact components, therefore the waveforms of the most relevant atoms are added with some impact components with damping coefficient ξ and frequency f increasing. In 116.67 h, when the fault is obvious, the waveforms of the most relevant atoms are mostly impact components with big ξ and f.

5.1.3. Fault pattern recognition

After extracting the atoms, we use frequencies of the 100 most relevant atoms and time-domain features as inputs of SVM to detect the weak impact components. The classification results are provided in Table 2 and Fig. 31.

5.1.4. Discussion

From Table 2 and Fig. 31, we can see that in 33.3–41.7, 66.7–75.0 and 116.7–125.0 test samples, that is when the roller bearing is normal or has serious failure, the proposed method and time-domain features all can recognize the condition with almost 100% accuracy. However, in 103.3–111.7 test samples, that is the fault incipient stage, time-domain features are unable to extract the weak fault signal, which test accuracy is 68.63%, while the proposed method still has 100% accuracy. Because in the fault incipient stage, the components of vibration signal changed but the amplitude of them did not change a lot. Therefore, the proposed method is able to detect the weak fault as the atoms that relevant to the original signal have changed a lot, but the time-domain features cannot.

5.2. Case 2

5.2.1. Experiment introduction

In this subsection, we use another bearing run-to-failure test to demonstrate the effectiveness of the proposed method. This test bearing is under constant load conditions to reflect the defect propagation processes. The test bench is composed of main body, control system, data acquisition computer system, power and drive system, hydraulic loading system and so on, as shown in Fig. 32.

The bench is designed as a supported beam structure as described in Fig. 33. The two steady bearings are N312 cylindrical roller bearings which are fixed at the middle of the shaft and the two test bearings are 30,311 tapered roller bearings which are installed on both ends of the shaft. The parameters of bearings are listed in Table 3. Axial load F_a is applied on bearing4 and transmitted to bearing 1 through the shaft. Radial load F_r is applied on the bushing and transmitted to the shaft through the steady bearings, finally applied on the test bearings. In this test, F_a is 15 kN and F_r is 27 kN. The rotation speed is 1500 r/min constantly.

The spindle motor speed and test load are controlled by computer. Vibration signals are exported through the screw which is connected to the outer race and collected by the Lance LC0401 High Sensitivity ICP accelerometer. YE6267 dynamic data collection and analysis system is used to record data. Sampling frequency is 10 kHz per channel and 32,768 points is collected every five minutes. As shown in Fig. 34. The temperatures are monitored by the thermocouple sensors.

In the end of the test, the bearing fails with a rolling element defects as shown in Fig. 35.

Four time-domain features of the same signal were used as inputs of SVM for comparison the same as Case 1. The waveforms of RMS are shown in Fig. 36.
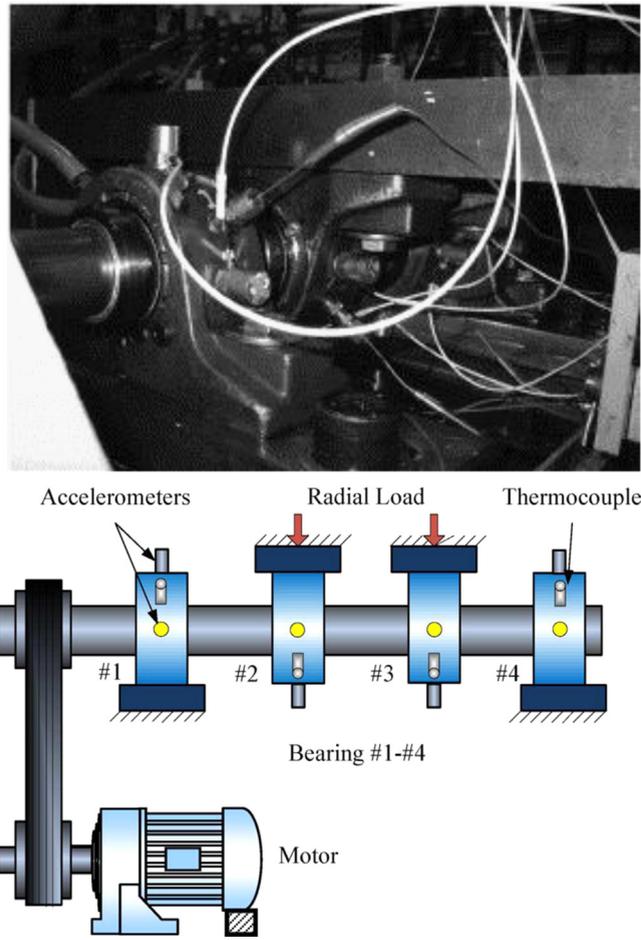
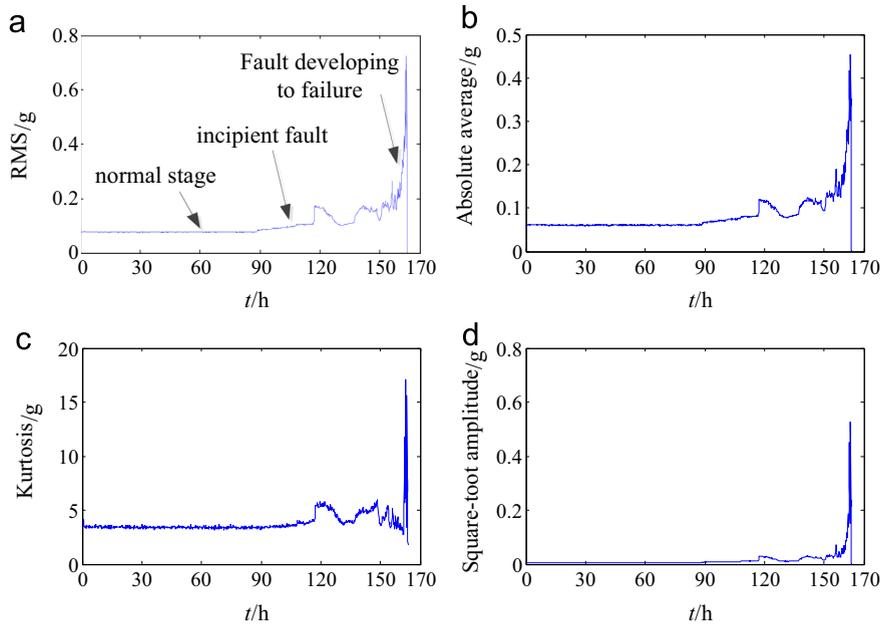**Fig. 25.** Bearing test rig.

Fig. 26. Waveforms of four time-domain features of failure bearing: (a) root means square (RMS); (b) absolute average; (c) kurtosis; (d) square-root amplitude.
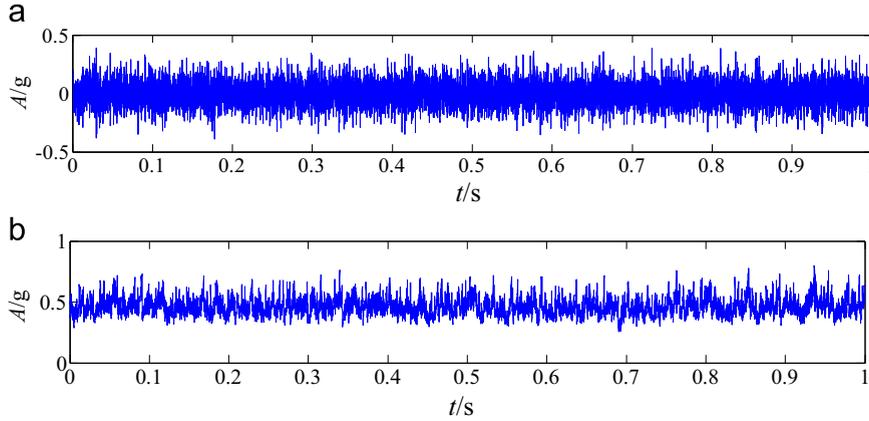
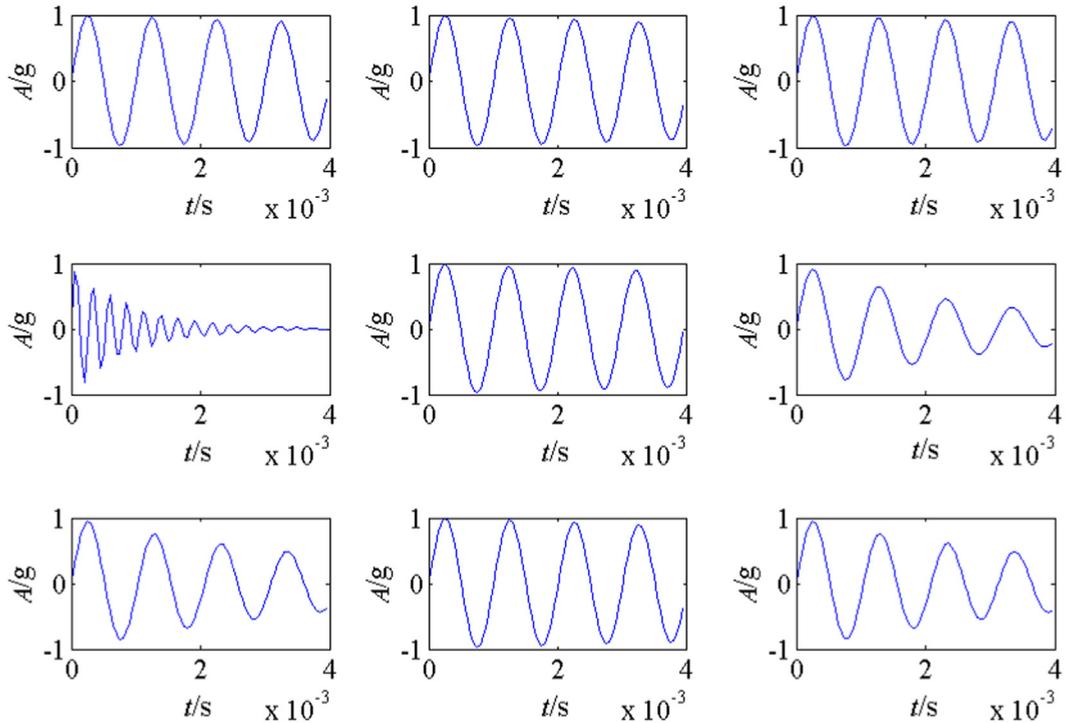**Fig. 27.** (a)Waveform of original signal; (b) waveform of correlation coefficient κ_γ.

Fig. 28. Waveforms of the 9 most relevant atoms in 18.33 h.

5.2.2. Features selection

After the test and the vibration information acquisition, short-time relevance matching algorithm (STRM) was utilized for feature extraction. The dictionary we used is $D = \left(g_\gamma(t)\right)_{\gamma \in \Gamma}$ which is constructed in 3.2 with parameters of damping coefficient ξ, frequency of the atom f_0 and the atom support width W_s. The parameter space Γ is determined as $\xi = \{0.005: 0.01: 0.2\} \cup \{0.2: 0.1: 0.9\}, f = \{500: 5: 5000\}, W_s = 0.004$, which is the same with case 1. In this experiment, we extract the first 100 biggest κ_γ and the corresponding atoms.

5.2.3. Fault pattern recognition

After extracting the atoms, we use frequencies of the 100 most relevant atoms as inputs of SVM to detect the weak impact components. Data from 58.33 h to 62.5 h and from 135.42 h to 139.58 h are used for training. The predict results based on SVM and different features are shown in Fig. 37. The classification accuracies are provided in Table 4 and Fig. 38.
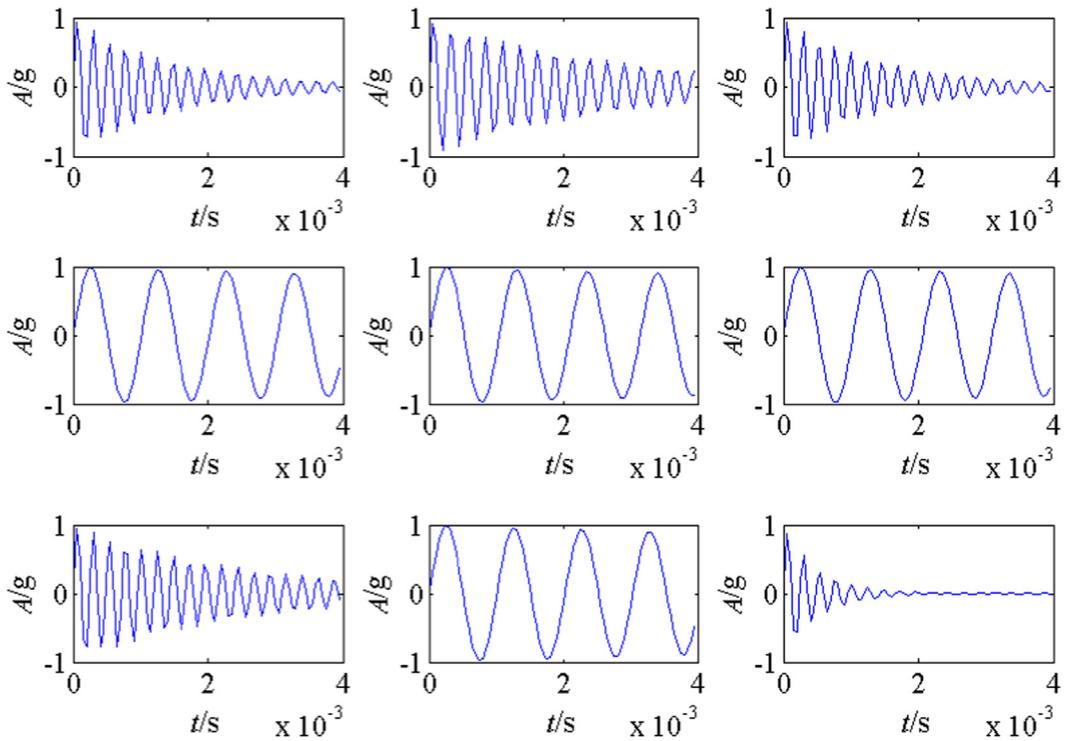
Fig. 29. Waveforms of the 9 most relevant atoms in 88.83 h.
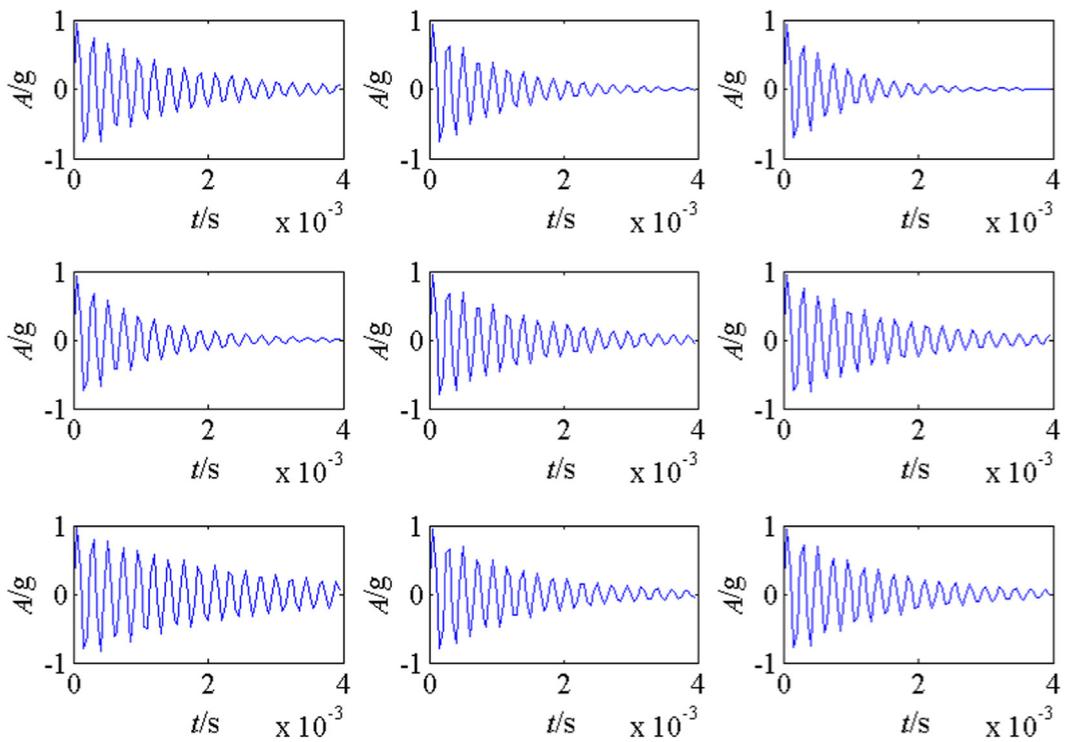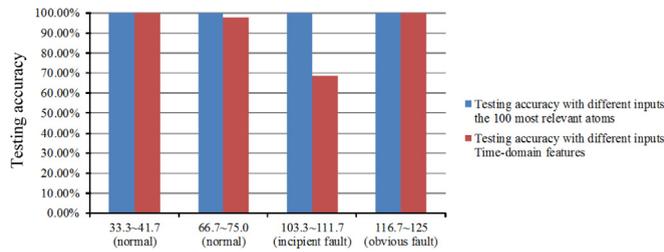

Fig. 30. Waveforms of the 9 most relevant atoms in 116.67 h.

Table 2
The diagnosis results with different features in SVM.

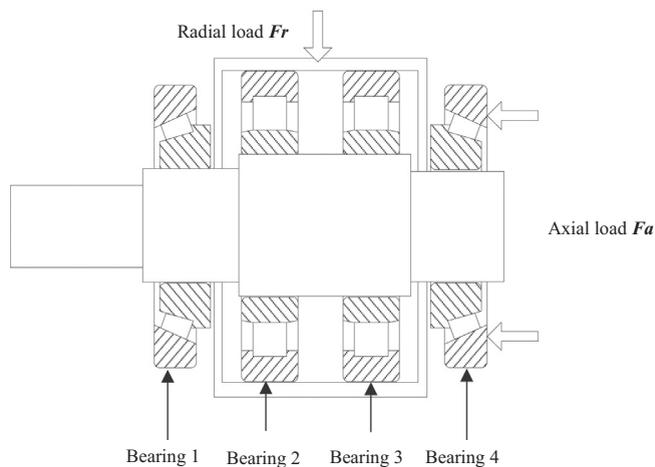| Test samples (time/h)	Testing accuracy with different inputs	
	The 100 most relevant atoms	Time-domain features
33.3–41.7 (Normal)	100.00%	100.00%
66.7–75.0 (Normal)	100.00%	98.04%
103.3–111.7 (Incipient fault)	100.00%	68.63%
116.7–125.0 (Obvious fault)	100.00%	100.00%

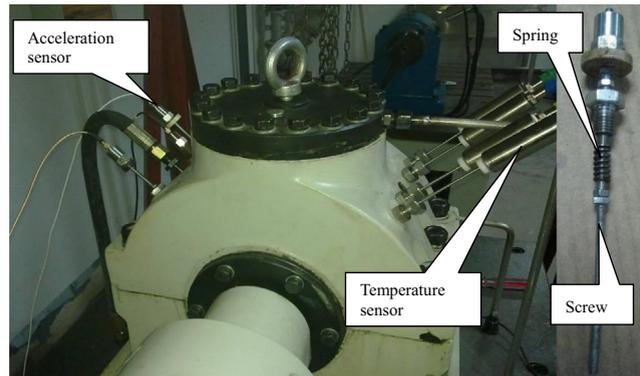**Fig. 31.** Testing accuracies with different features.
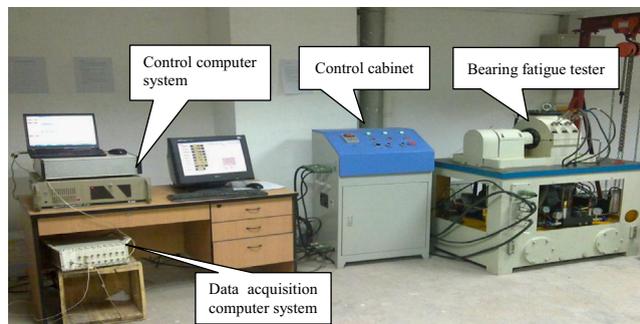
Fig. 32. Test bench of the bearing run-to-failure.

Fig. 33. The load diagram of four bearings.
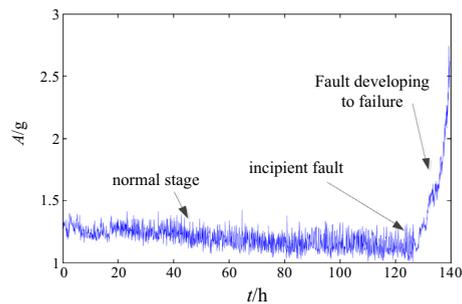
Table 3
Parameters of bearings.

Bearing type	Outer diameter (mm)	Inner diameter (mm)	Roller diameter (mm)	Roller number	Basic dynamic load (kN)
30,311	120	55	16.25	16	152
N312	130	60	19.1	16	212

Fig. 34. Location of accelerometer sensors and temperature sensors.

Fig. 35. The defects of the failure bearings.

Fig. 36. Waveform of root means square (RMS).

5.2.4. Discussion

From Fig. 37(a), we can see that if use time-domain features as inputs of SVM, the fault can be diagnose in 132 h, and after that there still are some samples be misdiagnosed. While in Fig. 37(b), when we use the parameters of atoms as inputs of SVM, the fault can be diagnose in 126.3 h, which is much earlier than traditional features. And there is no misdiagnosis after 126.3 h. What's more, according to the comparison result in Table 4 and Fig. 38, we can see in 126.17–130.33 and 129.50–133.67 test samples, that is the fault incipient stage, the initial fault cannot be distinguishable by SVM if use the traditional statistical indexes. While the proposed method is sensitive to the weak fault even if in the very early stage with 100% accuracy.
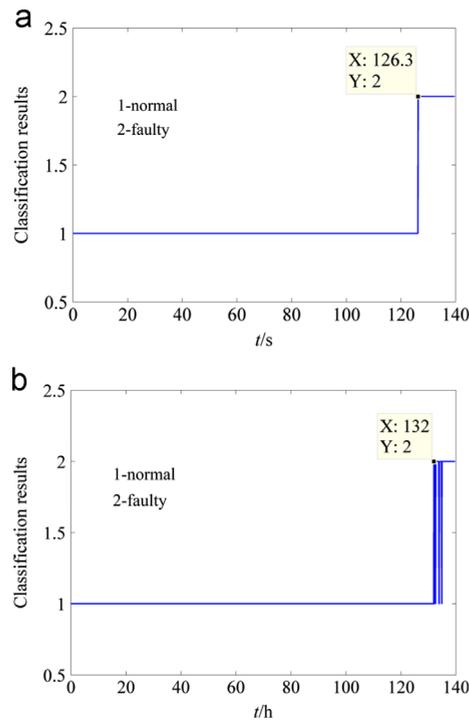
Fig. 37. SVM predict results based on the two different features: (a) based on the traditional time-domain features; (b) based on the parameters of most relevant atoms.

Table 4
The diagnosis results with different features in SVM.

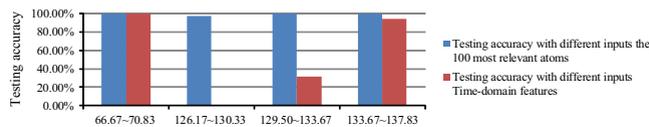| Test samples (time/h)	Testing accuracy with different inputs	
	The 100 most relevant atoms	Time-domain features
66.67–70.83 (Normal)	100.00%	100.00%
126.17–130.33 (Incipient fault)	97.06%	0.00%
129.50–133.67 (Incipient fault)	100.00%	31.37%
133.67–137.83 (Obvious fault)	100.00%	94.12%

Fig. 38. Testing accuracies with different features.

6. Conclusions

In this paper, we construct a dictionary based on the mechanism of roller bearings, and propose a novel fault diagnose method based on short-time matching and SVM, to overcome the limitations of traditional sparse representation and fault diagnose methods. Firstly, in order to reduce the calculation cost, we use a short-time signal $f(\tau, \tau+W_s)$ instead of the entire signal $x(t)$ to match the dictionary D. Secondly; a new correlation measurement index based on correlation filter and MP was proposed to avoid the influence by l^2-norm in conventional methods. Thirdly, in order to make the method more suitable for failure diagnosis, we extract the atoms which were mostly relevant to the original signal as inputs to SVM instead of representing signal as a linear combination of atoms. The simulation and comparison results illustrate that the proposed method behaves better than MP, EMD, SK and traditional time-domain methods in detecting the weak impact signal with strong noise condition. The applications of roller bearing run-to-failure test show that the proposed method can reflect the running state of rolling bearing in real time and detect the fault in the embryonic effectively and efficiently. In addition, this method could be applied to the fault diagnose of other key machine components and more fields such as feature extraction and noise reduction.

Acknowledgments

References

[1] Y. Yang, D. Yu, J. Cheng, A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM, Measurement 40 (2007) 943–950.
[2] B. Li, M.Y. Chow, Y. Tipsuwan, J.C. Hung, Neural-network-based motor rolling bearing fault diagnosis, IEEE Trans. Ind. Electron. 47 (2000) 1060–1069.
[3] C. Sun, Z.S. Zhang, Z.J. He, Z.J. Shen, B.Q. Chen, W.R. Xiao, Novel method for bearing performance degradation assessment-A kernel locality preserving projection-based approach, Proc. Inst. Mech. Eng. C J. Mech. 228 (2014) 548–560.
[4] J. Yuan, Z.J. He, Y.Y. Zi, Y.G. Lei, Z. Li, Adaptive multiwavelets via two-scale similarity transforms for rotating machinery fault diagnosis, Mech. Syst. Signal Process. 23 (2009) 1490–1508.
[5] J. Yuan, Z.J. He, Y.Y. Zi, Gear fault detection using customized multiwavelet lifting schemes, Mech. Syst. Signal Process. 24 (2010) 1509–1528.
[6] R.J. Hao, Z.K. Peng, Z.P. Feng, F.L. Chu, Application of support vector machine based on pattern spectrum entropy in fault diagnostics of rolling element bearings, Meas. Sci. Technol. 22 (2011).
[7] Y. Yang, J.S. Cheng, K. Zhang, An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems, Measurement 45 (2012) 561–570.
[8] N.E. Huang, Z. Shen, S.R. Long, M.L.C. Wu, H.H. Shih, Q.N. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. A Math. Phys. 454 (1998) 903–995.
[9] R.Q. Yan, R.X. Gao, X.F. Chen, Wavelets for fault diagnosis of rotary machines: a review with applications, Signal Process. 96 (2014) 1–15.
[10] S.B. Wang, W.G. Huang, Z.K. Zhu, Transient modeling and parameter identification based on wavelet and correlation filtering for rotating machine fault diagnosis, Mech. Syst. Signal Process. 25 (2011) 1299–1320.
[11] Z.K. Zhu, R.Q. Yan, L.H. Luo, Z.H. Feng, F.R. Kong, Detection of signal transients based on wavelet and statistics for machine fault diagnosis, Mech. Syst. Signal Process. 23 (2009) 1076–1097.
[12] Y.G. Lei, J. Lin, Z.J. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mech. Syst. Signal Process. 35 (2013) 108–126.
[13] B.Q. Chen, Z.S. Zhang, C. Sun, B. Li, Y.Y. Zi, Z.J. He, Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors, Mech. Syst. Signal Process. 33 (2012) 275–298.
[14] R.Q. Yan, R.X. Gao, Rotary machine health diagnosis based on empirical mode decomposition, J. Vib. Acoust. 130 (2008).
[15] J.S. Cheng, D.J. Yu, J.S. Tang, Y. Yang, Local rub-impact fault diagnosis of the rotor systems based on EMD, Mech. Mach. Theory 44 (2009) 784–791.
[16] Y.X. Wang, Z.J. He, Y.Y. Zi, A comparative study on the local mean decomposition and empirical mode decomposition and their applications to rotating machinery health diagnosis, J. Vib. Acoust. 132 (2010).
[17] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, Mech. Syst. Signal Process. 21 (2007) 108–124.
[18] S.G. Mallat, Z.F. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Trans. Signal Process. 41 (1993) 3397–3415.
[19] J. Lin, Feature extraction of machine sound using wavelet and its application in fault diagnosis, NDT&E Int. 34 (2001) 25–30.
[20] W.X. Lai, P.W. Tse, G.C. Zhang, T.L. Shi, Classification of gear faults using cumulants and the radial basis function network, Mech. Syst. Signal Process. 18 (2004) 381–389.
[21] B. Samanta, Artificial neural networks and genetic algorithms for gear fault detection, Mech. Syst. Signal Process. 18 (2004) 1273–1282.
[22] K. Grochenig, Describing functions-atomic decompositions versus frames, Monatsh. Math. 112 (1991) 1–41.
[23] S.S.B. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (2001) 129–159.
[24] G. Davis, S. Mallat, M. Avellaneda, Adaptive greedy approximations, Constr. Approx. 13 (1997) 57–98.
[25] A. Widodo, B.S. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mech. Syst. Signal Process. 21 (2007) 2560–2574.
[26] Z.J. Shen, X.F. Cheng, Z.J. He, Hybrid intelligent fault diagnosis based on adaptive lifting wavelet and multi-class support vector machine, Int. C Wavel Anal Pat (2010) 415–420. 10.1109/ICWAPE.2010.5576405.
[27] H.Y. Yang, J. Mathew, L. Ma, Fault diagnosis of rolling element bearings using basis pursuit, Mech. Syst. Signal Process. 19 (2005) 341–356.
[28] L.C. Freudinger, R. Lind, M.J. Brenner, Correlation filtering of modal dynamics using the Laplace wavelet, Proc. Soc. Photo Opt. Instrum. 3243 (1998) 868–877.
[29] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics, J. Sound. Vib. 289 (2006) 1066–1090.