Mechanical Systems and Signal Processing 102 (2018) 346–363

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Gear fault diagnosis based on the structured sparsity time-frequency analysis

Ruobin Sun [a,b], Zhibo Yang [a,b], Xuefeng Chen [a,b,*], Shaohua Tian [a,b], Yong Xie [c]

[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
[c] School of Aerospace, Xi'an Jiaotong University, Xi'an, PR China

ARTICLE INFO

Article history:
Received 15 June 2017
Received in revised form 10 September 2017
Accepted 16 September 2017

Keywords:
Structured sparse time-frequency analysis
Vibration components separation
Periodic impulsive vibration extraction
Gear fault diagnosis

ABSTRACT

Over the last decade, sparse representation has become a powerful paradigm in mechanical fault diagnosis due to its excellent capability and the high flexibility for complex signal description. The structured sparsity time-frequency analysis (SSTFA) is a novel signal processing method, which utilizes mixed-norm priors on time-frequency coefficients to obtain a fine match for the structure of signals. In order to extract the transient feature from gear vibration signals, a gear fault diagnosis method based on SSTFA is proposed in this work. The steady modulation components and impulsive components of the defective gear vibration signals can be extracted simultaneously by choosing different time-frequency neighborhood and generalized thresholding operators. Besides, the time-frequency distribution with high resolution is obtained by piling different components in the same diagram. The diagnostic conclusion can be made according to the envelope spectrum of the impulsive components or by the periodicity of impulses. The effectiveness of the method is verified by numerical simulations, and the vibration signals registered from a gearbox fault simulator and a wind turbine. To validate the efficiency of the presented methodology, comparisons are made among some state-of-the-art vibration separation methods and the traditional time-frequency analysis methods. The comparisons show that the proposed method possesses advantages in separating feature signals under strong noise and accounting for the inner time-frequency structure of the gear vibration signals.

1. Introduction

Gears are one of the most important components in mechanical transmission systems. It is widely used in machining, transportation, aerospace, wind generation and other fields [1–4]. Gears, however, usually cause mechanical shutdown even casualties due to its rugged working environment. Therefore, the condition monitoring and fault diagnosis for gears are of great significance in ensuring the operational safety of systems. Vibration signal analysis is a common and effective technique for gear damage detection. The vibration signals contain rich sources of information from the various mechanical components. By extracting fault characteristics of the gears embedded in these signals, the health condition and fault type can be confirmed. However, gear transmission systems are elastic mechanical systems which generate dynamic response under

* Corresponding author at: The State Key Laboratory for Manufacturing Systems Engineering, Xi'an, PR China.
E-mail address: chenxf@mail.xjtu.edu.cn (X. Chen).

https://doi.org/10.1016/j.ymssp.2017.09.028
0888-3270/

dynamic excitation. The coupling between mechanical components and inevitable noise interference further compounds the vibration signals.

In order to extract gear damage information from vibration signals effectively, a large amount of intensive and fruitful research has been accomplished, such as the shaft synchronous signal averaging methods [5–7], the empirical mode decomposition methods [8], wavelets transform methods [9]. And to improve the performance of extracting gear damage information from vibration signals, sparsity methodology is introduced in gear vibration signals modeling. The sparse model is that signals can be linearly represented as a few atom signals on a redundant basis. The rationality of using the sparse model to machinery diagnosis can be explained by the following: Firstly, from the spectrum or the time-frequency distribution of vibration signals, the presentation coefficients are sparse indeed. Moreover, it has been confirmed that mammalian vision systems and auditory systems have similar sensing principle as the sparse model [10,11], and thus the experienced maintainers are able to diagnose their familiar machines just by listening to their working sounds. These facts reveal that the sparse model is suitable for machinery fault diagnosis.

The sparse model is widely used in fault diagnosis in recent years. The greatest strength of the sparse model is to improve the capability on describing arbitrary complex signals, and the related signal decomposition method is called sparse representation (SR). For instance, Fourier analysis provides a poor representation of time localized signals, and wavelet analysis is not well-adapted to represent high-frequency signals in a narrow bandwidth [12–14] These shortcomings are attributed to the attempt to represent arbitrary signals with a limited set of basis functions in a fixed form. As a contrast, the SR decomposed the signals over redundant and over-completed dictionaries, thus being free from the limitation. So far, the SR methods have been widely studied for gear damage detection. Chen et al. [15] extracted gear impulsive components based on the sparse dictionary learning and redundant representations over the learned dictionary. Fan et al. [16] proposed a transient feature extraction technique based on the sparse representation using the wavelets basis. Cui et al. [17] employed the composite dictionary and multi-atom matching decomposition and reconstruction algorithm to extract characteristics of gear fault signals. Feng et al. [18] used the shift invariant K-means singular value decomposition dictionary learning method to analyze planetary fault signals and diagnose the localized and distributed gear faults successfully. He et al. [19] used the matching pursuit and correlation filtering to separate the coupling components of gearbox vibration signals under intensive background noise.

In the above researches, the sparsity of coefficients was enforced by optimization equations with l_1 norm penalty. The main drawback of the l_1 norm penalty is that all the coefficients are treated independently, thus a natural extension is the structured sparsity which can be achieved by mixed norm penalty. Structured sparsity has been successfully applied in many engineering fields, such as computer vision [20], text processing [21], audio processing [22], but there are only few reports about the application of the structured sparsity in mechanical fault diagnosis. Most mechanical vibration signals are highly structured in time domain and frequency domain. Local defects, i.e. pitting, tooth root crack and tooth break, are the most common mechanical faults. When the local defect exists, periodic impulses will be produced during the rotation of the gear. Therefore the vibration signals will consist of three main components: steady modulation, periodic impulses, and noise. Both the steady modulation and the periodic impulses contain the fault features. Extracting these feature is the key issue in gear fault diagnosis.

The structured sparse time-frequency analysis (SSTFA) method is a structured sparsity method proposed by Kai to process audio signals [23]. The SSTFA uses the mixed-norm priors on time-frequency coefficients and a weighted generalized thresholding operator to lead in the structured sparse representation of signals. The SSTFA method enforces sparsity in one domain, while simultaneously, having diversity and persistence in the other domain. The SSTFA can better account for the inner structure of the defective gear vibration signals. By aid of different structured time-frequency priors, different components of the defective gear vibration can be separated, and a more accurate diagnosis conclusion can be given.

In this paper, the SSTFA is utilized to analyze the mechanical vibration signals and recognize the gear damage. In Section 2, the concept of structured sparsity is briefly introduced. Then based on the vibration signal model of the locally defective gear, the process of the SSTFA for defective gear detection is proposed. In Section 3, a numerical simulation analysis is carried out and some comparisons with the state-of-the-art methods are conducted. In Section 4, a fault simulation experiment is carried out and an engineering application of wind turbine gear fault diagnosis is illustrated. The results show that the SSTFA method improves the accuracy of the gear damage detection. Finally, some conclusions are drawn in Section 5.

2. Structured sparsity time-frequency analysis

The main idea of the sparse representation is to represent signals with the over-completed redundant function sets which are called dictionaries instead of orthogonal bases. It is a more flexible way to model arbitrary signals encountered in engineering applications.

2.1. Sparse regularization

The vibration signals are often distorted by strong noises, which are caused by sensor imperfection, poor running environment, communication errors, and so on [24]. The observed vibration signal with additive noise can be written as:

$$\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{y}$ is the observed signal, $\boldsymbol{f}$ is the target signal, and $\boldsymbol{e}$ is the error.

In engineering applications, the signals are all discrete due to sampling. The aim of our work is to approximate the recovery of the signal $\boldsymbol{f} \in \mathrm{R}^N$. Consider the sparse model of the signal, one can represent the signal $\boldsymbol{y}$ by basis functions in the dictionary as: $\boldsymbol{f} = \sum_{\gamma\in\Gamma} c_\gamma \boldsymbol{\varphi}_\gamma$ (Γ is the corresponding index set and the coefficients $\boldsymbol{c} \in \mathrm{C}^K$). $\boldsymbol{\Phi}:\ \mathrm{R}^N \to \mathrm{C}^K$ is known as the synthesis operator with $\boldsymbol{\Phi} = (\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_\gamma, \ldots)$. In above model, the coefficients are unknown. To obtain a fine approximation for them, the discrepancy should be minimized:

$$\Delta(\boldsymbol{c}) := \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi c}\|_2^2 \tag{2}$$

This is a linear inverse problem, and it does not have a unique solution. Thus the regularization method has been widely adopted to solve this problem. By adding some applicable or reasonable constraints on the coefficients in form of a penalty measure $\Psi:\ \mathrm{C}^K \to \mathrm{R}_0^+$, one can obtain the regularized functional

$$\mathcal{L}(\boldsymbol{c}) := \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi c}\|_2^2 + \lambda\Psi(\boldsymbol{c}) \tag{3}$$

and seek $\hat{\boldsymbol{c}}$ yields

$$\hat{\boldsymbol{c}} = \arg\ \min_{\boldsymbol{c}} \mathcal{L}(\boldsymbol{c}) \tag{4}$$

The value $\lambda > 0$ is named as sparsity level since it provides the weight for the penalty term. The larger λ, the harsher penalty will be taken into account, and vice versa [25].

The simple but powerful idea of sparse representation is to approximate the signals by a few atoms $\boldsymbol{\varphi}_\gamma$ as the trial basis. The most natural way is to choose l_0 -penalty measure $\Psi(\boldsymbol{c}) = \|\boldsymbol{c}\|_0$, where $\|\cdot\|_0$, denotes the number of non-zero coefficients in $\boldsymbol{c}$. Minimizing such a penalty leads to an NP-hard problem, which is usually relaxed into a l_1 norm convex penalty by choosing $\Psi(\boldsymbol{c}) = \|\boldsymbol{c}\|_1$ [26].

2.2. Mixed norms and generalized thresholding operator

One of the main limitations of the l_1 norm sparse model is that all the coefficients are treated independently. Most mechanical vibration signals are highly structured, for example, the impulsive components of signals are existing in some certain time and located in some frequency band in the time-frequency distribution. Therefore the structures of a signal corresponding to the physical prior could be used for its processing [27]. In order to take the advantage of time-frequency structures of the gear vibration signal, we consider the dictionary as the Gabor frame which is an over-completed time-frequency dictionary and replace the l_1 norm penalty by the mixed norm $l_{p,q}$ which acts as different roles in groups (indexed by g in the sequel, can be either time or frequency) and their members (indexed by m) [23].

Gabor frames consist of a set of atoms $\boldsymbol{\varphi}_{k,j} = \boldsymbol{M}_{bj}\boldsymbol{T}_{ka}\boldsymbol{\varphi}$, where $\boldsymbol{T}_x$ and $\boldsymbol{M}_\omega$ denote the time- and frequency-shift-operator respectively. $\boldsymbol{\varphi}$ is a standard window function. a and b are the time and frequency sampling constant, and $j = 0, \ldots, J-1,\ k = 0, \ldots, K-1$, with $Ka = Jb = L$. Accordingly, the expansion coefficients of the Gabor frame dictionary are denoted by the two-dimensional variable c_{kj}. The notation (k,j) refers to the time-frequency indices of the Gabor expansion.

The mixed norm $l^{p,q}$ with the group-member structure yields:

$$\Psi(c) = \|c\|_{p,q} = \left(\sum_g \left(\sum_m |c_{g,m}|^p\right)^{q/p}\right)^{1/q} \tag{5}$$

Specifically, the $l^{2,1}$ ($p = 2,\ q = 1$) penalty is known as the Group-Lasso (GL), which promotes sparsity in groups and diversity in members. In the case of $p = 1,\ q = 2$, the $l^{1,2}$ penalty is termed Elitist-Lasso (EL) which promotes sparsity in members and diversity in groups [28].

The iteration soft-thresholding algorithm (ISTA), which solves Eq. (4) with the l_1 norm penalty [29], also yields a solution to the generalized minimization problem induced by Eq. (5). Kai [30] replaced the standard soft thresholding by a generalized thresholding operator and has taken the fast ISTA(FISTA) methods into consideration as presented in [31]. Kai's method accelerates the convergence velocity of the iterative procedure. The generalized thresholding operator is defined as $\mathbb{S}_{\lambda,\xi}(\mathbb{z}_{\mathbb{g},\mathbb{m}}) = \mathbb{z}_{\mathbb{g},\mathbb{m}}(\mathbb{1} - \xi(\mathbb{z}))^+$, where $b \in \mathrm{R}$, $b^+ := \max(b, 0)$, ξ is the thresholding function, which is defined as:

$$p = 1, q = 1:\ \xi^L(c_{g,m}) = \frac{\lambda}{|c_{g,m}|} \quad (Lasso) \tag{6}$$

$$p = 2,\ q = 1:\ \xi^{GL}(c_{g,m}) = \frac{\lambda}{\left(\sum_m |c_{g,m}|^2\right)^{\frac{1}{2}}} \quad (GL) \tag{7}$$

$$p=1,\ q=2:\ \xi^{EL}(c_{g,m})=\frac{\lambda}{1+M_g\lambda}\frac{\|\boldsymbol{c}_g\|_1}{|c_{g,m}|}\quad (EL) \tag{8}$$

where $\boldsymbol{c}_g=(c'_{g,1},\ldots,c'_{g,M_g})$ and $\{c'_{g,m'}\}_{m'}$ denotes for each group g the sequence of scalars $|c_{g,m}|$ in descendant order. M_g denotes some natural numbers related to the magnitudes of coefficients in the group $(c_{g,1},\ldots,c_{g,M})$ [32]. The solution to Eq. (5) is then given by the FISTA as follows.

2.3. Time-frequency neighborhood

In some vibration signal processing issue, the groups of GL and EL cannot be identified explicitly. Therefore the time-frequency neighborhood is defined as follows to cover a wider spectrum of vibration signals:

For an index $\gamma=(g,m)$ in a structured index set $\mathcal{I}$, a weighted neighborhood $N(\gamma)=\{\gamma'\in\mathcal{I}:\ w_\gamma(\gamma')\neq 0\}$ with the weights w_γ defined on $\mathcal{I}$ such that $w_\gamma(\gamma')\geqslant 0$ for all $\sum_{\gamma'\in N(\gamma)}w_\gamma(\gamma')^2=1$. The coefficients $c_{g,m}$ undergo shrinkage according to the energy of the time-frequency neighborhood, and the neighborhoods are more flexible to the groups of GL and EL using weighting and overlap. Then with $\eta_N(c_\gamma)=\left(\sum_{\gamma'\in N(\gamma)}w_\gamma(\gamma')^2|c_{\gamma'}|^2\right)^{1/2}$, the thresholding functions of the generalized thresholding operator is further extended by the following convolutions:

$$\xi^{WGL}=\xi^{L}*\eta_N\quad (windowed\ GL\ (WGL)) \tag{9}$$

$$\xi^{PEL}=\xi^{EL}*\eta_N\quad (persistent\ EL\ (PEL)) \tag{10}$$

$$\xi^{PGL}=\xi^{GL}*\eta_N\quad (persistent\ GL\ (PGL)) \tag{11}$$

2.4. The process of gear fault analysis

The vibration signals are basically induced by the gear meshing and the rotation of gear shafts. The presence of a local fault will generate the impulse force when the gear is engaged in meshing. The impulse force will cause the variation of vibration amplitude and phase which generates steady modulation phenomenon [33]. As the consequence of the steady modulation, vibration signals exhibit sidebands around the meshing frequency and its harmonics. The traditional gear fault diagnosis methods mainly focus on the detection of sideband structure in spectrums. This is not the best approach for fault diagnosis as the manufacturing and assembling for intact gears also induce modulation [34]. Furthermore, in the complex gear transmission systems, such as planetary gear train, sidebands may not be obvious even in the seriously defective gear. This can be interpreted by the long vibration transmission path, which involves strong noise interference related to the rotation of the planet carrier. Compared with modulation phenomenon, the periodic impulse feature of vibration is more robust for local damage detection in gear. Thus, combining sideband structures with periodic impulse features is a promising way to detect gear damage.

For the stationary gear vibration signal, i.e. rotation speed of the machine does not change with time, the harmonic components of the shaft rotation and steady modulation components are sparse in frequency domain but persistent in time domain, as shown in Fig. 1(a). Since the time-frequency neighborhood consists of a fixed frequency index and some succes-

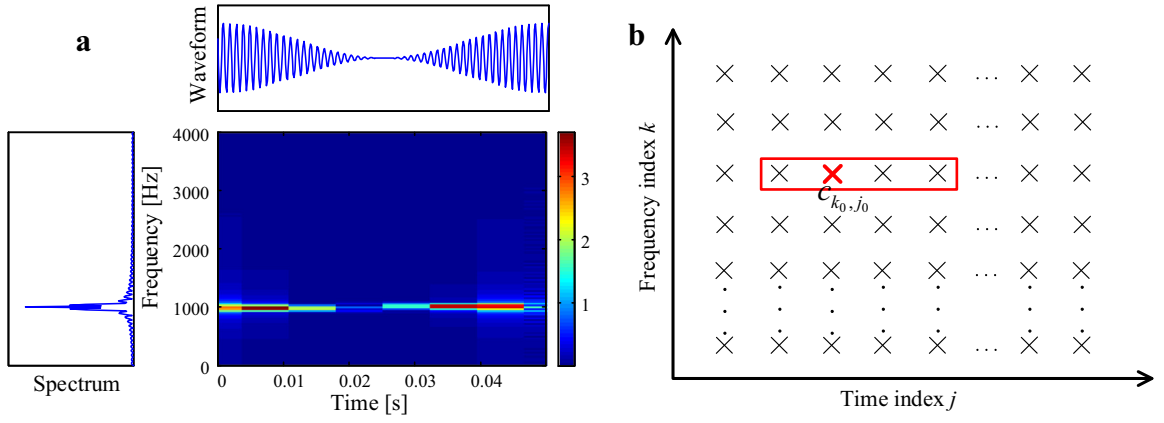

Fig. 1. (a) The time-frequency distribution of the steady modulation signal. (b) The sketch of time-frequency neighborhood of the WGL for the stationary harmonic and modulation components.

sive time indexes, the WGL can be used to exploit the time-frequency priors of the steady modulation components. As illustrated in Fig. 1(b), the neighborhood of the index $(k_0,\ j_0)$ is $\{(k_0-1,\ j_0),\ (k_0,\ j_0),\ (k_0+1,\ j_0),\ (k_0+2,\ j_0)\}$.

However, the periodic impulse signals induced by local defect are sparse in time domain but persistent in frequency domain, as shown in Fig. 2(a). Furthermore, the inspected impulse is damped, namely it is persistent in time domain. In this situation, the PGL is a good selection to introduce this kind of time-frequency priories. Groups are chosen as the time indexes, and members are chosen as the frequency indexes. Some additional persistence in frequency is also added, such as using neighborhoods which contain two successive time indexes, as illustrated in Fig. 2(b).

The length of persistent time indexes can be determined by the following heuristic method. Firstly, identify the modal dynamic parameters of the impulsive vibration by the correlation filtering using the Laplace wavelet [35]. The Laplace wavelet is formulated as an impulse response of a single mode system, as follows:

$$\psi(f,\zeta,\tau,t)=\psi_\gamma(t)=\begin{cases}Ae^{-2\pi f\zeta(t-\tau)}e^{-j2\pi f\sqrt{1-\zeta^2}(t-\tau)}, & t\in[\tau,\tau+W_s]\\ 0 & \text{otherwise}\end{cases} \tag{12}$$

where $\gamma=\{f,\zeta,\tau\}$ is the modal parameters, and W_s is the support length of the wavelet. Discretize the parameter and find the maximum correlation coefficient at the every τ moment, as formulated by Eq. (13).

$$\kappa(\tau)=\max_{f,\zeta}\kappa_\gamma^\tau=\max_{f,\zeta}\frac{|\langle\psi_\gamma^\tau(t),x(t)\rangle|}{\|\psi_\gamma\|_2\|x\|_2} \tag{13}$$

Find the peak of the curve $(\tau,\ \kappa(\tau))$, and the corresponding parameters $(\bar{f},\ \bar{\zeta})$ are the modal parameters of the vibration. Thus the envelope of the impact response can be identified. Take the end of the impulsive signal where the amplitude drops to 0.01 of the maximum amplitude, the persistent length can be estimated by:

$$n=\left\lceil\left(\frac{-f_s\ln(0.01)}{2\pi\bar{f}\bar{\zeta}}-l_w\right)\Big/\Delta l_w\right\rceil+1 \tag{14}$$

where f_s is the sampling frequency, l_w is the window length of the Gabor frame, and Δl_w is the shift length of the window. The above process is illustrated by Fig. 3.

The different components of the gear vibration signal can be extracted respectively by the different thresholding functions of the generalized thresholding operators. So the SSTFA procedure for defective gear detection is described as follows:

STEP 1: Select an appropriate analysis frequency band. In application, the sampling frequency is often much higher than twice of the analyzing frequency to guarantee that the high-frequency components are reasonably sampled. In fact, the analysis frequency contains the first three orders of the gear meshing frequency and the main frequency range of the impulse response is wide enough for gear fault diagnosis. The pre-processing of the signal using a low-pass filter to choose an appropriate frequency band is highly recommended. The selected filter can reduce the data number to improve calculation speed and get a more readable time-frequency distribution.
STEP 2: Choose a relatively long window of the Gabor frame. The time-frequency neighborhood consists of a fixed frequency index and some successive time indexes. Then use the WGL thresholding function and solve the sparse regularization inverse problem to extract harmonic and steady modulation components.
STEP 3: Subtract the harmonic component from the signal to get the residual signal. Choose a relatively short window of the Gabor frame, the time indexes as groups and neighborhood of two successive time indexes. Then use the PGL to extract the impulsive component from the residual signal.

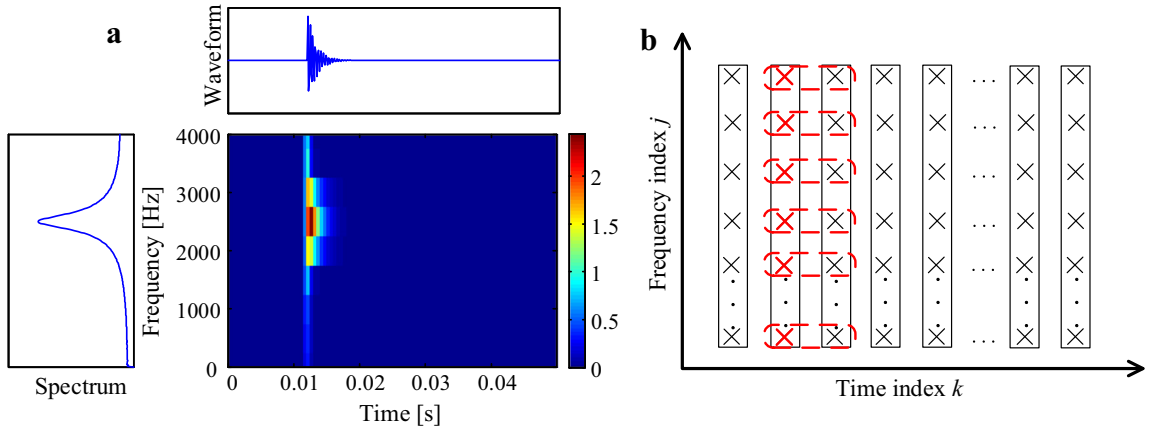

Fig. 2. (a) The time-frequency distribution of the impulse signal. (b) The sketch of time-frequency neighborhood of the PGL for impulsive components: groups are defined by time indexes as shown in the rectangles and persistent of two-time indexes is introduced as shown in the dashed.

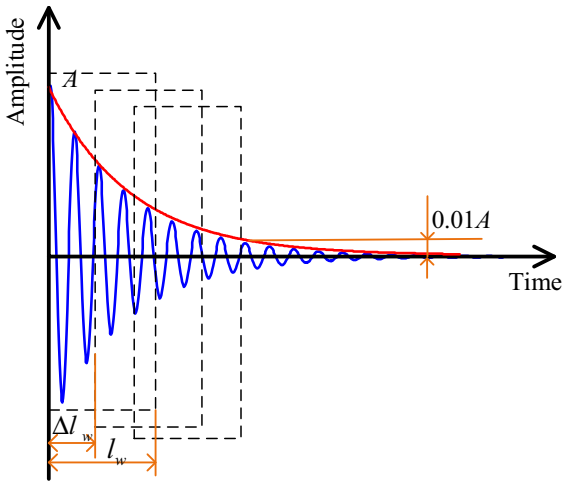

Fig. 3. The sketch of how to choose the persistent length.

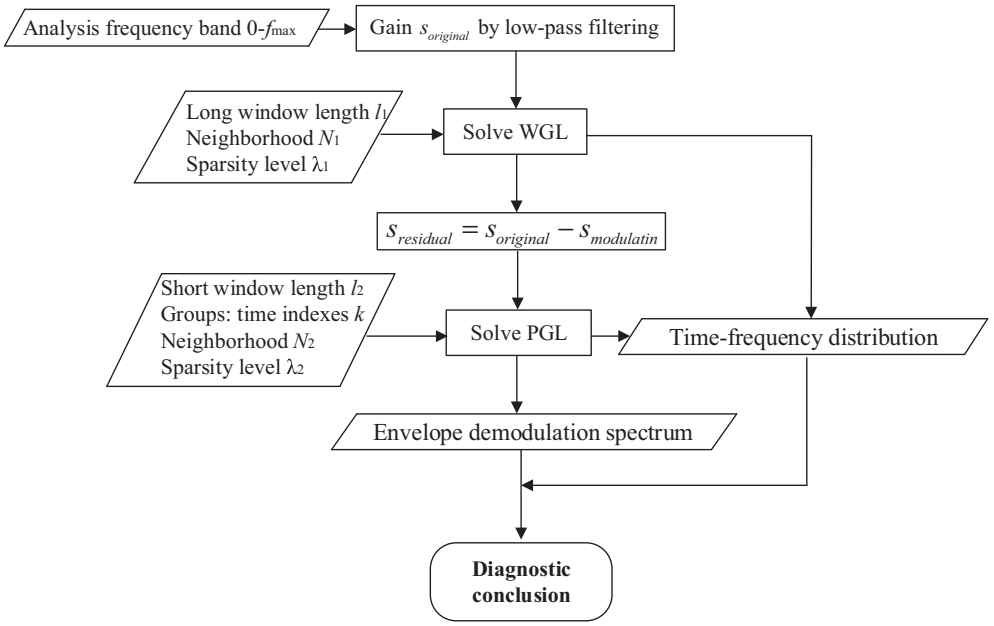

Fig. 4. The process of the SSTFA for the damage detection of gear.

STEP 4: Plot the time-frequency distribution of the steady modulation components and the impulsive components in the same diagram respectively. According to the time-frequency distribution or the envelope demodulation spectrum of the impulsive components, a diagnostic conclusion is obtained.

The process of the SSTFA for defective gear detection is presented in Fig. 4.

3. Numerical simulation and comparison

Based on above analysis, there are three main components in the defective gear vibration signals including the harmonic and steady modulation components $s_h(t)$ (induced by gear meshing and rotation of the shaft), the periodic impulsive components $s_i(t)$ (generated from gear local damage), and other additive noise $n(t)$ .Without loss of generality, the initial phase of the harmonic signals is neglected. The observed vibration signal of defective gears can be described as:

$$
\begin{aligned}
s(t) &= s_h(t) + s_i(t) + n(t) \\
s_h(t) &= \sum_m A_{sm}\cos(2\pi m f_{r1} t) + \sum_n B_{sn}\cos(2\pi n f_{r2} t) + [1 + A_h\cos(2\pi f_{r1} t)]\cdot\cos(2\pi f_z t + B_h\cos(2\pi f_{r1} t)) \\
s_i(t) &= \sum_k A_i \exp[-\zeta\omega_n(t - kT)]\cdot\sin[\omega_n\sqrt{1-\zeta^2}(t - kT)]
\end{aligned}
\tag{15}
$$

where $\sum_m A_{sm}\cos(2\pi m f_{r1} t)$ and $\sum_n B_{sn}\cos(2\pi n f_{r2} t)$ are the rotational vibration components and their higher orders of the two gear shafts. To simplify the signal, only the first order of the rotation frequency is considered. A_{s1} is set to be 1 and B_{s1} is set to be 0.5. f_{r1} is 20 Hz and f_{r2} is 40 Hz. The meshing frequency f_z is 1000 Hz. Assume that a local fault exists on shaft 1, the amplitudes of amplitude modulation (AM) and frequency modulation (FM) A_h and B_h are both set to be 1. $s_i(t)$ is the single side attenuation impact simulation signal, $A_i = 2.5$ is amplitude factor of the impact signal, ζ is the damping ratio equaling to 0.05, ω_n is the undamped natural frequency equaling to $2500/2\pi$ rad $\cdot$ s^{-1} and T is the period of impact signal equaling to $1/f_r$. $n(t)$ is white Gaussian noise with zero mean and standard deviation $\sigma = 0.5$ .The sampling frequency of the signal is set as 8 kHz, and the sampling number is 4096. The waveforms of the synthetic signals are shown in Fig. 5.

3.1. Parameter analysis

In order to evaluate the effect of different parameters on the algorithm, various parameters are employed to process the above signal. The parameters of the comparison include: the window length l of Gabor frame and time-frequency neighborhood length n in both the WGL and the PGL procedures. The result is evaluated by the mean squared errors (MSE) of the separated components and original signals. To be as fair as possible, the sparsity level λ is optimized respectively by the trial and error approach to gain the best separation result. The MSE curves that vary with neighborhood length in different window length are presented in Fig. 6.

The MSE1 of the steady modulation components decreases with the increase of neighborhood length when window length becomes larger. That is due to the persistence of the simulated steady modulation components. Choosing a longer neighborhood should be suitable for this situation, but the MSE1 is not influenced by the neighborhood length when the length is too small. That can be attributed to the fact that the small window length is not suitable for extracting steady modulation components. To separate impulsive components from the inspected signal, the optimal length of the neighborhood length could be calculated by Eq. (14). The identified modal parameters are $\bar{f} = 2530$ Hz and $\bar{\zeta} = 0.048$, which is close to the simulated results. The optimal lengths of the neighborhood are 4, 2, 1, and 1, which are respectively corresponding to the window lengths 20, 30, 40 and 50. For the MSE2 curves of the impulsive components, the minimum value just happens when the length is close to the optimal neighborhood length. Furthermore, the MSE2 increases with the window length. It shows

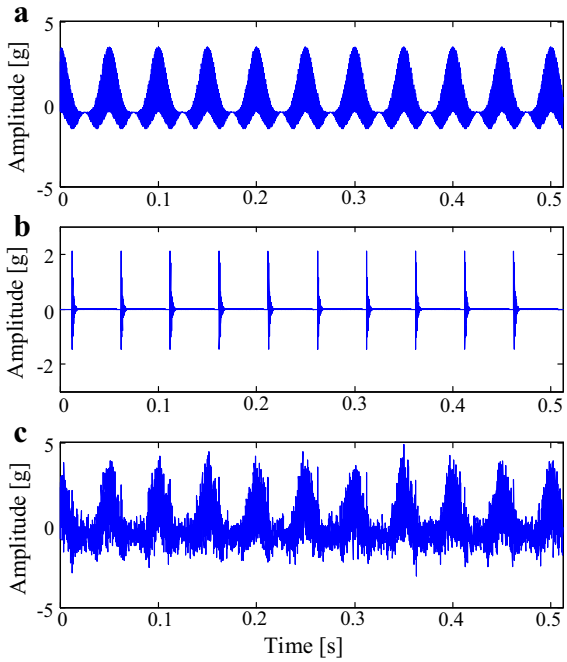

Fig. 5. The simulation signals: (a) the steady modulation and harmonic component; (b) the periodic impulsive component; (c) the synthetic signal with white noise.

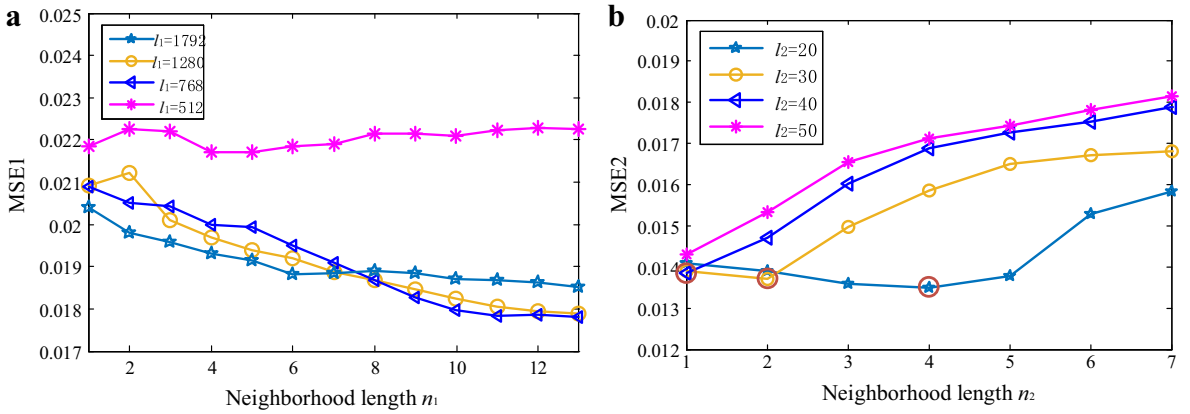

Fig. 6. The MSE curves vary with neighborhood length in different window lengths. (a) MSE curve of the steady modulation components; (b) MSE curve of the impulsive components. The red circle is the minimum of the curve. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that choosing a smaller window length and optimal neighborhood length could gain a better separation result for impulsive components. Thus, we choose the window length $l_1 = 1024$, the neighborhood length $n_1 = 13$ for the WGL procedure, and the window length $l_2 = 32$, and the corresponding optimal neighborhood length n_2 for the PGL procedure in this paper.

3.2. Method comparison

The simulated signal is then processed by the SSTFA. The reconstructed steady modulation components and periodic impulsive components are displayed in Fig. 7. It can be observed from Fig. 7 that the stationary components and the impulsive components are completely separated except that each end of the signal is corrupted due to the boundary effect. As shown in the partial enlarged detail in Fig. 8, it is seen that the reconstructed component has a high quality, which is very close to the original signals.

Finally, the time-frequency distribution of the steady modulation and impulsive component are drawn in the same time-frequency diagram as shown in Fig. 9(a). The time-frequency distribution diagram has the lower background noise and higher resolution in both time domain and frequency domain. It is interpreted by the good denoising performance of the method and the appreciable separation of different time-frequency structures. The low-frequency part is the rotational frequency of the two shafts. The periodicity of the impulse signals is 0.05 s. One can see that the impact energy is concentrated around the damped natural frequency $\omega_n\sqrt{1-\zeta^2} = 2487.5$ Hz. From the partial enlarged detail in Fig. 9(b), the interval of the sideband around meshing frequency is 20 Hz, which is also the modulated rotational frequency of the shaft. To sum up, the time-frequency structure of the simulation signal is fully revealed.

For comparison, the same synthetic signal is analyzed by an SR decomposition method, morphological component analysis (MCA) [36], and a classic signal self-adaptive decomposition method, empirical mode decomposition (EMD) [37] respec-

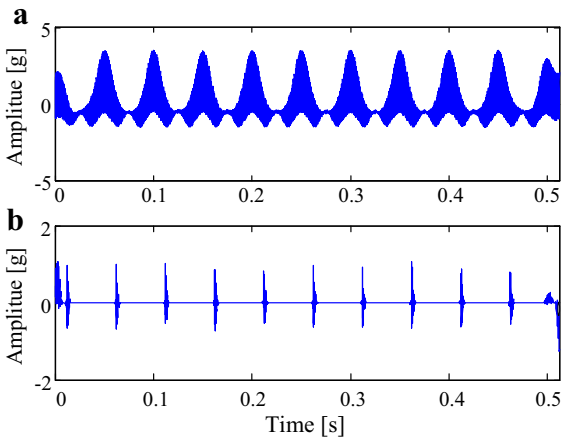

Fig. 7. The reconstructed signal of (a) the steady modulation components; and (b) the periodic impulsive components.

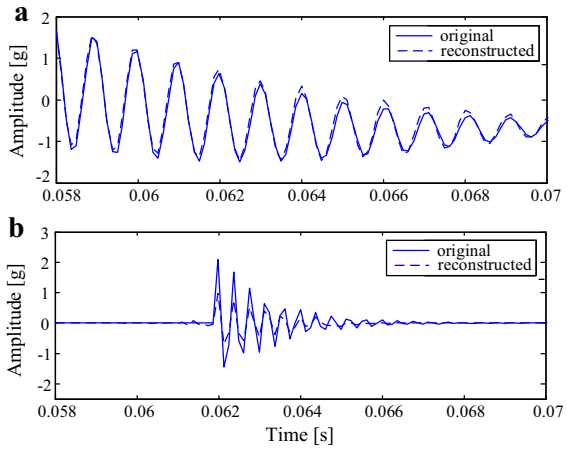

Fig. 8. The partial enlarged detail of the original and reconstructed signal of (a) the steady modulation components; and (b) the periodic impulsive component.

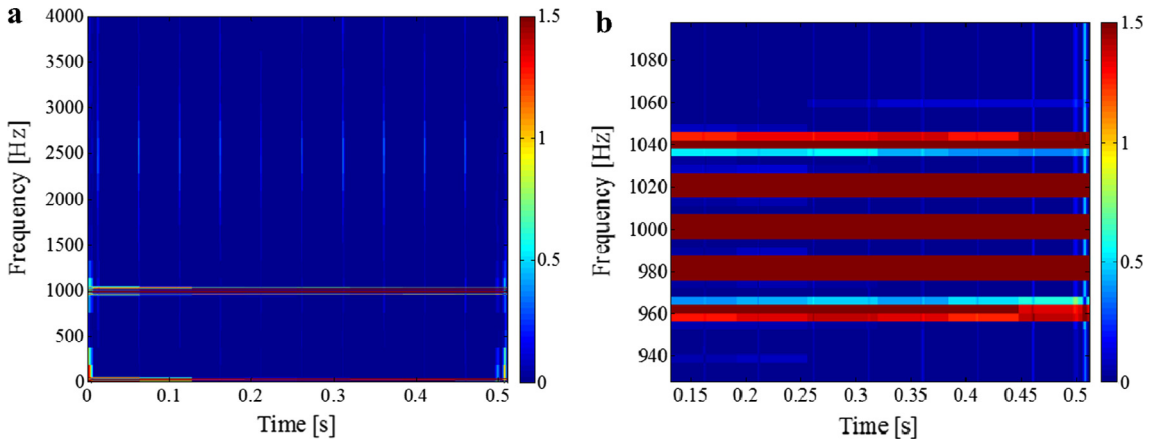

Fig. 9. (a) The structured sparse time-frequency distribution of the steady modulation and the impulsive component, and (b) its partial enlarged detail.

tively. Firstly, the MCA is used to decompose $s(t)$. The MCA sparsely represents different signal components in the different dictionaries. The steady modulation components are sparsely represented by the discrete cosine dictionary and the impulsive components are sparsely represented by the db8 (Daubechies wavelets with 8-order vanishing moments) wavelets dictionary. The SR problems are solved by orthogonal matching pursuit (OMP) method respectively. The results are shown in Fig. 10. The impulsive components are not only corrupted by endpoint effect but also interfered by harmonic components. To further evaluate the performance of the different methods, the MSE of the separated components and original signals are listed in Table 1, which indicates that the SSTFA has the higher decomposition accuracy than the MCA. For the vibration signals in this experiment (on a corei7-4790 @3.6 GHz computer), the running time of the SSTFA is 1.025 s, while the running time of the MCA is 16.142 s. Therefore, the SSTFA also performs better than the MCA on efficiency, which makes the SSTFA a suitable choice for in-situ applications (see Table 2).

Secondly, the EMD is employed in decomposition the simulation signal, and the first six intrinsic mode functions (IMFs) are shown in Fig. 11. It can be seen that the steady modulation and impulsive components are not separated completely. The IMF1 contains part of the impulsive components. Besides, IMF3 mainly includes the impulsive components, but it is still interfered by the harmonic components. The comparison demonstrates that the SSTFA has the more powerful capability of separating different components than the MCA and the EMD.

Finally, the simulation signal is analyzed by means of a traditional time-frequency analysis method, short time Fourier transform (STFT), and a sparse time-frequency representation method. The Hamming window with the length of 128 is used in the STFT. The sparse time-frequency representation method obtains the sparse representation on the db8 wavelets dictionary utilizing the basis pursuit denoising and visualizes the coefficients in the time-frequency plane. The time-frequency diagrams of the signal are illustrated in Fig. 12. The STFT method is corrupted by noise. Furthermore, due to the restriction of the Heisenberg uncertainty principle, the time localization and frequency resolution cannot present their

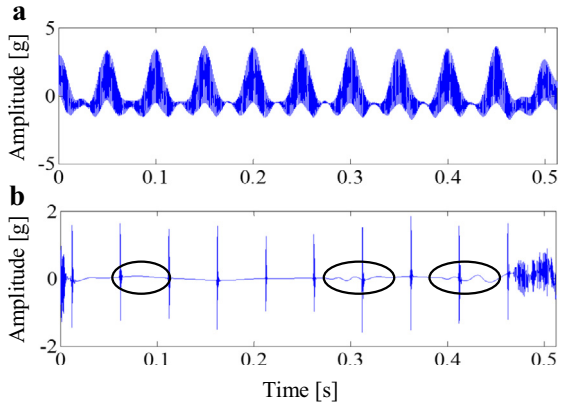

Fig. 10. The decomposition results of the MCA: (a) the steady modulation components and (b) the impulsive components.

Table 1
Fast iteration soft-thresholding algorithm.

Algorithm (FISTA)
$\mathbb{S} = \mathbb{S}_{\lambda,\xi}$
choosing arbitrary $\boldsymbol{c}^0 = \boldsymbol{b}^1$
$t_1 = 1$
repeat
$\boldsymbol{c}^n = \mathbb{S}(\boldsymbol{b}^n + \boldsymbol{\Phi}^*(\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{b}^n))$
$t_{n+1} = \frac{1}{2}\left(1 + \sqrt{1 + 4t_n^2}\right)$
$\boldsymbol{b}^{n+1} = \boldsymbol{c}^n + \left(\frac{t_{n-1}}{t_{n+1}}\right)(\boldsymbol{c}^n - \boldsymbol{c}^{n-1})$
Until convergence

Table 2
The MSE obtained by the SSTFA and MCA.

	Steady modulation components error $e_1(t)$	Impulsive components error $e_2(t)$
SSTFA	0.0227	0.0125
MCA	0.0235	0.0170

best performance simultaneously. The periodic impulses can be vaguely distinguished, but the sidebands of the steady modulation components are hardly identified. The sparse representation method gains a high frequency resolution. However, due to the structure priori information of the time-frequency coefficients has not been considered, the sideband information cannot be well-recognized either. In this sense, their performances are inferior to that of the SSTFA method for the investigated problem.

4. Experimental verification and engineering application

To verify the effectiveness of the SSTFA, a fault simulation test of detecting a damaged planet gear in planetary gear train is performed. Thereafter, the algorithm is applied in the damage detection of wind turbine driving chain gear.

4.1. Fault simulation test

The experiment was conducted on an SQI gearbox fault simulator as shown in Fig. 13. The test rig contains a two-stage planetary gearbox and a two-stage fixed-axis gearbox. The driving train is powered by a 3-hp motor and has a magnetic brake for loading. The kinematic diagram of the gearboxes is illustrated in Fig. 14. A simulated broken tooth fault is conducted on one of the planet gears in the first planetary stage as shown in Fig. 15. The vibration signals are acquired by an accelerometer implemented on the planetary gearbox near the first stage ring gear. The signals are recorded by the CoCo80 data acquisition instrument with the sampling frequency 10240 Hz, and the signals are finally displayed and analyzed in a laptop.

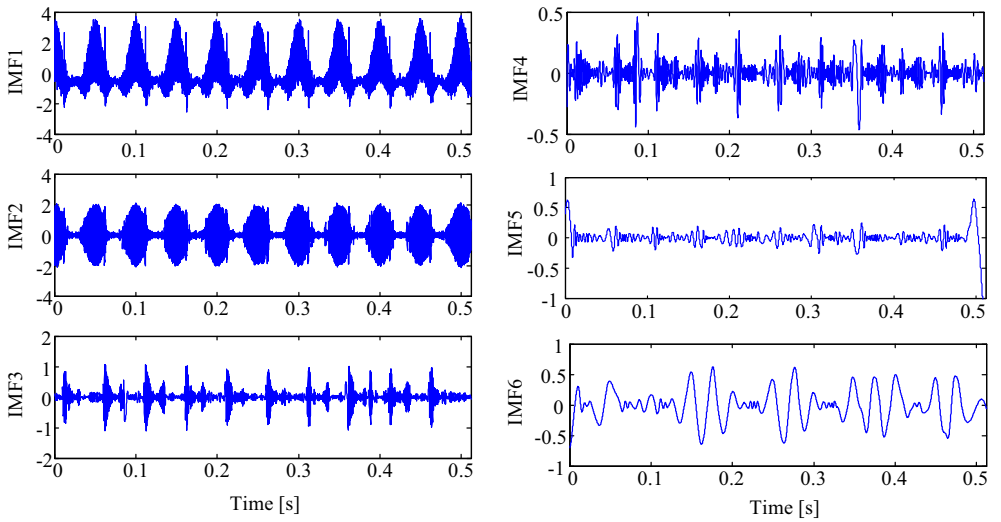

Fig. 11. The EMD analysis results of the vibration signal in the simulation.

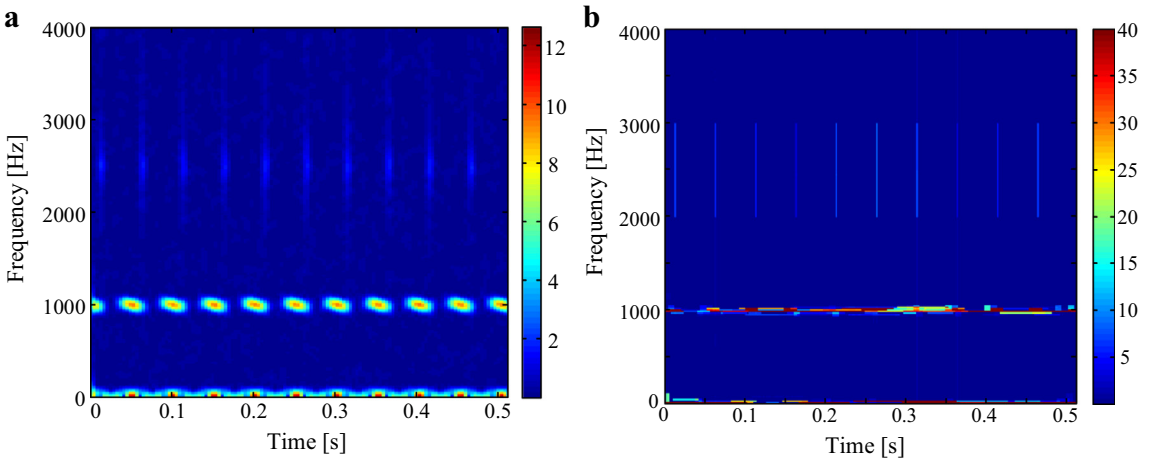

Fig. 12. (a) The time-frequency diagram of the simulation signal using the STFT method. (b) The time-frequency diagram of the simulation signal using basis pursuit denoising on the db8 wavelets dictionary.

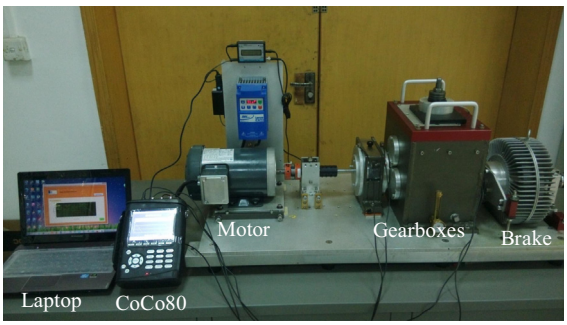

Fig. 13. The SQI gearbox fault simulator and the measurement system.

The input speed of the motor is 1800 rpm. The load torque is 10.88 N m. Under these conditions, the input torque is 1.98 N m. The meshing frequency is 500 Hz and the local fault characteristic frequency of the first stage planet gear is 25 Hz, which can be calculated via Eqs. (16) and (17).

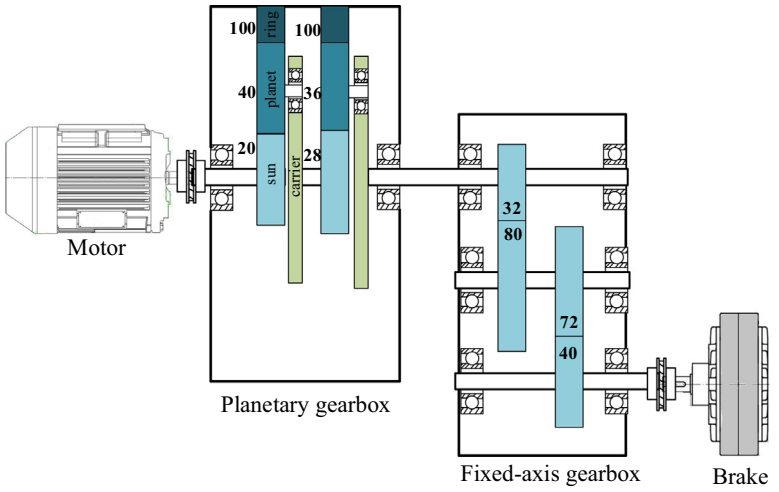

Fig. 14. The kinematic diagram of the gearboxes. The numbers are the tooth number of the corresponding gear.

Fig. 15. Broken tooth fault of the first stage planet gear.

$$f_m = f_i \frac{Z_s Z_r}{(Z_s + Z_r)} \tag{16}$$

$$f_p = 2\frac{f_m}{Z_p} = 2f_i \frac{Z_s Z_r}{Z_p(Z_s + Z_r)} \tag{17}$$

where Z_s, Z_p, Z_r are the numbers of the teeth of the sun gear, planet gear, and ring gear, respectively. f_m, and f_i refer to the meshing frequency and input rotational frequency, respectively.

A segment of the signals is intercepted and the analysis frequency band is chosen as 0–5120 Hz. The waveform and spectrum are shown in Figs. 16 and 17, respectively. Impulsive components cannot be clearly identified in the original waveform. In the low frequency region of the spectrum, the first four order of the input shaft rotational frequency can be seen clearly in Fig. 17(b), which is due to the unbalance of the shaft system caused by the fault planetary gear. Moreover, from the locally enlarged detail in Fig. 17(c), there is no obvious phenomenon of fault characteristic frequency modulating the meshing frequency. In the envelope demodulation spectrum (Fig. 18), the fault characteristic frequency is completely submerged in the interference caused by strong noise in the original signals. Hence, the further analysis using the SSTFA is necessary.

For extraction the steady modulation components, the window length of the Gabor frame l_1 is chosen as 1024 and the neighborhood length n_1 is chosen as 20. For extraction the impulsive component, the window length l_2 is chosen as 32 and the neighborhood length n_2 is chosen as 3, which is estimated by Eq. (14). The sparse level $\lambda_1 = 0.15$ and $\lambda_2 = 0.06$, which are optimized by the trial and error approach on another segment of the signals using the MSE evaluation criteria. The length of Fig. 19 illustrated the separated different components, and Fig. 20 is the envelope demodulation spectrum of the impulsive component. From the envelope demodulation spectrum, the fault characteristic frequency can be identified. The structured time-frequency distribution of the signal is shown in Fig. 21, from which the period of the impulsive component can be identified clearly. The period is 0.04 s, which corresponds to the local fault feature frequency of the first stage planet gear. Furthermore, the frequency band of the impulse can be observed from the time-frequency distribution. It can be

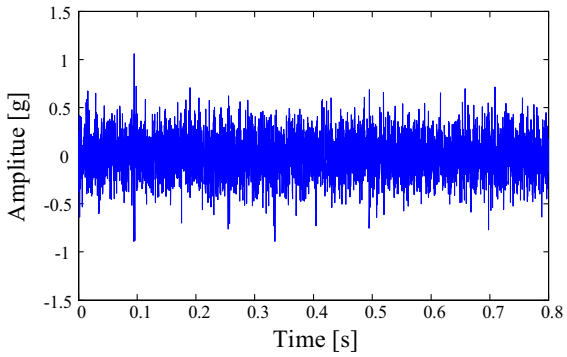

Fig. 16. A segment of the experimental vibration waveform.

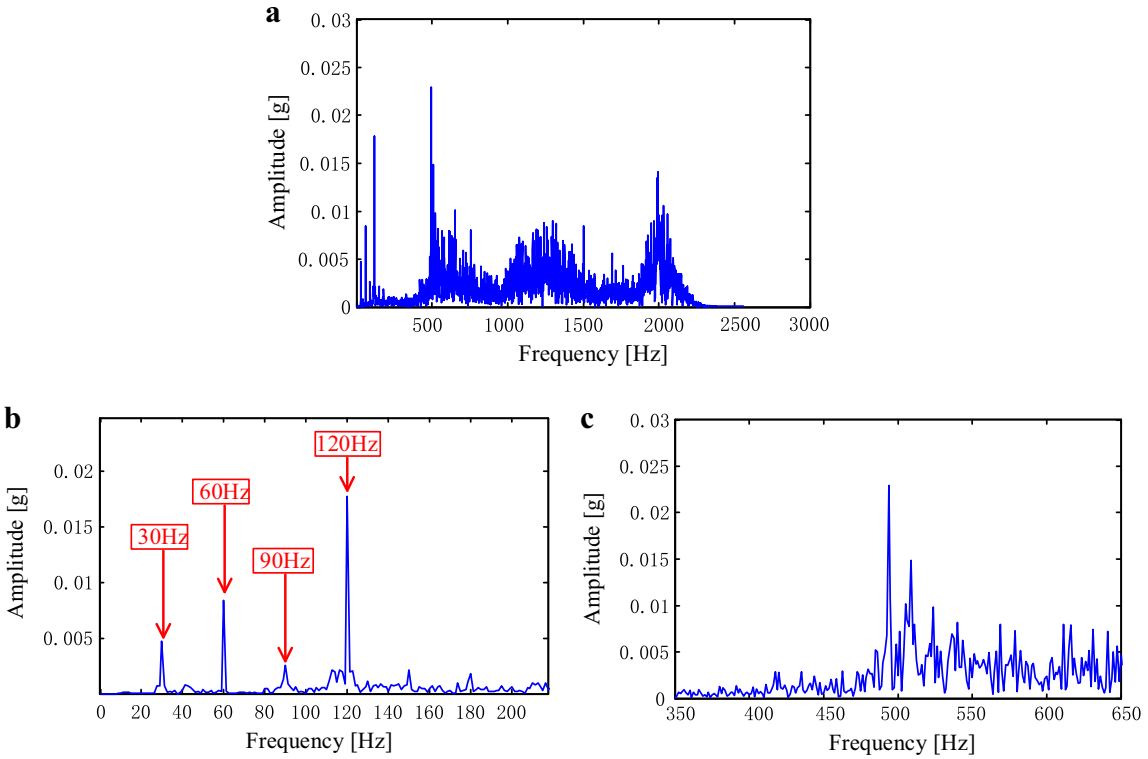

Fig. 17. (a) The spectrum of the signal. (b) The locally enlarged detail of the spectrum around the input rotational frequency and its harmonics. (c) The locally enlarged detail of the spectrum around the first stage meshing frequency.

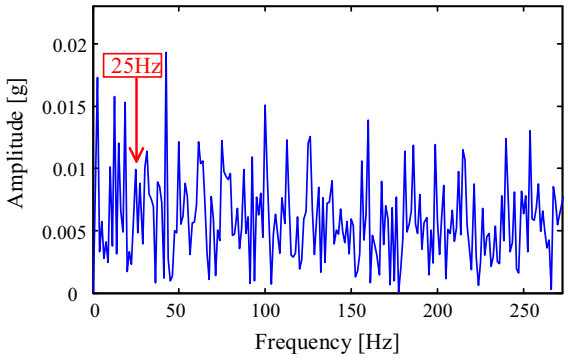

Fig. 18. The envelope demodulation spectrum of the inspected signal.

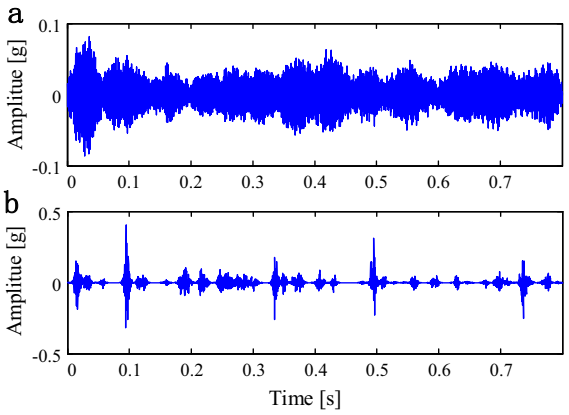

Fig. 19. The separated steady modulation components and the impulsive components.

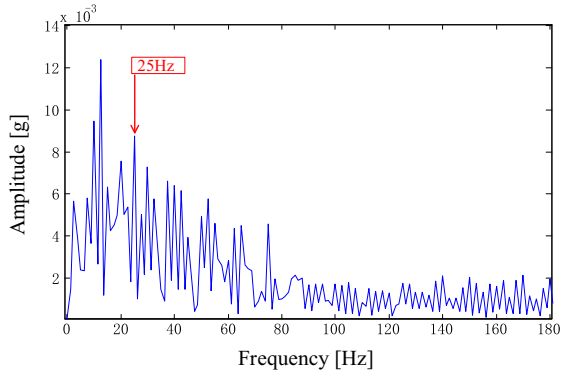

Fig. 20. The envelope demodulation spectrum of the impulsive component.

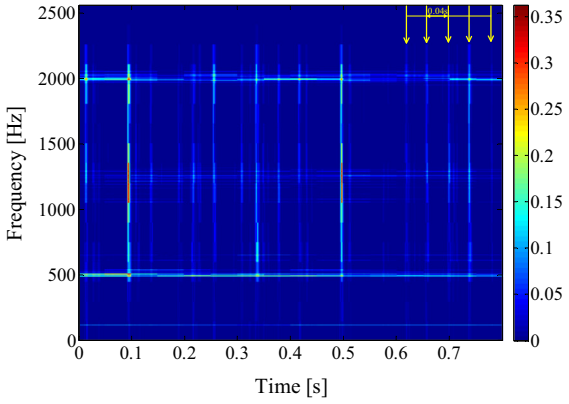

Fig. 21. The structured sparse time-frequency distribution of the inspected signal.

inferred that the impulse energy in the range of 500–2000 Hz are caused by resonance oscillations of the gear system when the locally damaged tooth meshes with the others.

4.2. Wind field test

During one inspection of a wind field, an alarm occurred in a wind turbine of the condition monitoring system (CMS). The installed capacity of the wind turbine is 2 MW. The root mean square (RMS) value of the gearbox vibration signals in the wind turbine was over the standard. The alarm accelerator is implemented on the bearing cover of the high-speed shaft

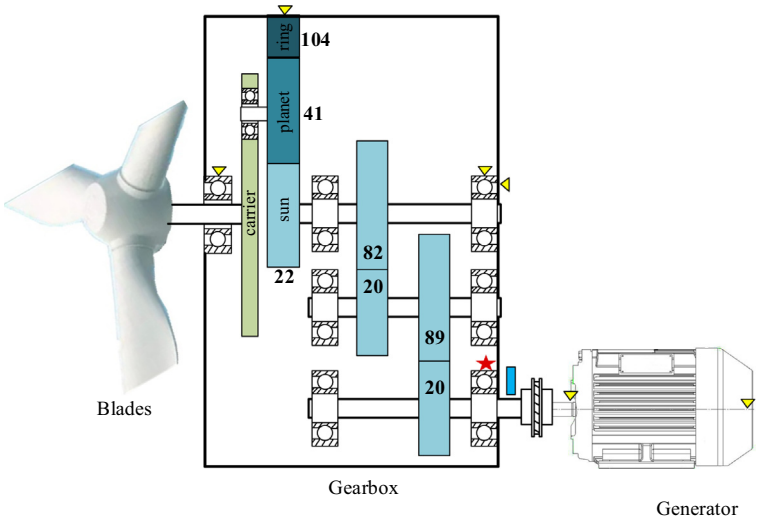

Fig. 22. The drivetrain kinematic diagram of the wind turbine. The numbers are the tooth number of gears. The accelerometer location of the CMS is represented by triangles, and the alarm accelerometer is represented by a star. The tachometer is marked by a rectangle.

Table 3
The characteristic frequencies of the fixed-axis gearbox at 1795.8 rpm (units: Hz).

	Meshing frequency	Gear shaft frequency	Pinion shaft frequency
Low speed stage	134.47	1.64	6.73
High speed stage	598.57	6.73	29.92

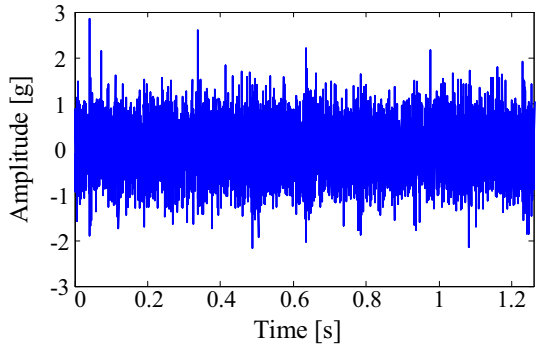

Fig. 23. One segment of the inspected vibration signal.

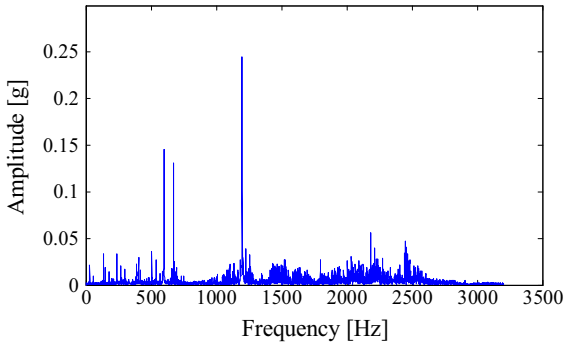

Fig. 24. The frequency spectrum of the wind turbine signal.

of the gearbox and a tachometer is implemented on the high-speed shaft. The drivetrain kinematic diagram and the sensor layout are shown in Fig. 22. The sampling frequency is set as 25,600 Hz.

From the tachometer records, the average rotational speed of the high-speed shaft is obtained as 1795.8 rpm. According to the kinematic relation, the characteristic frequency of the fixed-axis gearbox is listed in Table 3. Firstly, the decimation filter is used to achieve a decimation factor 8, i.e. the analysis frequency band is chosen as 0–3200 Hz. One segment of

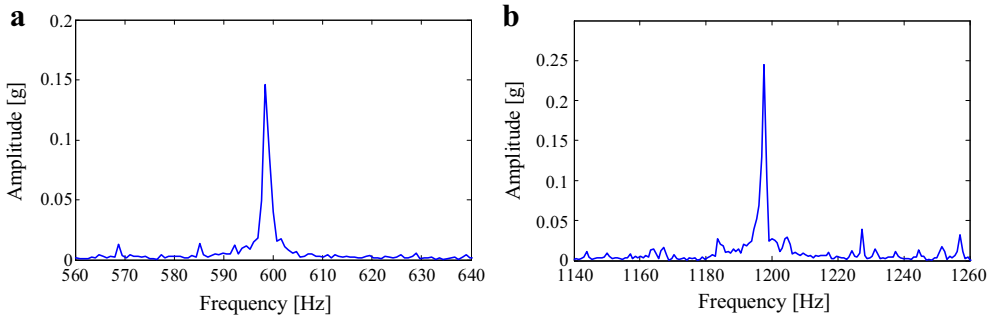

Fig. 25. The locally enlarged detail of the spectrum shown in Fig. 24. (a) The frequency band around the first order of the meshing frequency. (b) The frequency band around the second order of the meshing frequency.

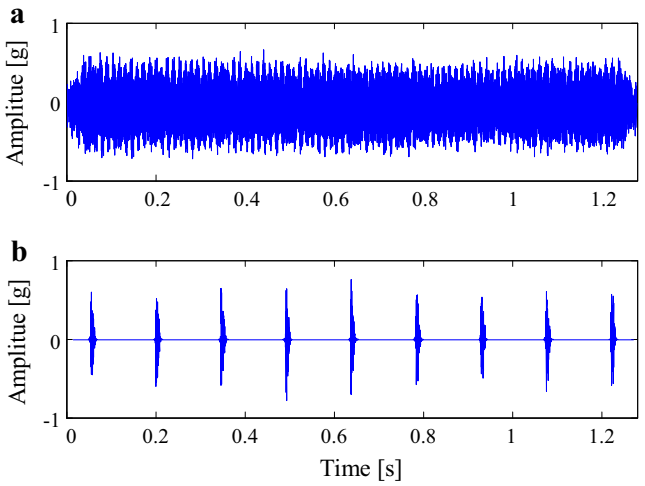

Fig. 26. The separated components of the steady modulation components and impulsive components.

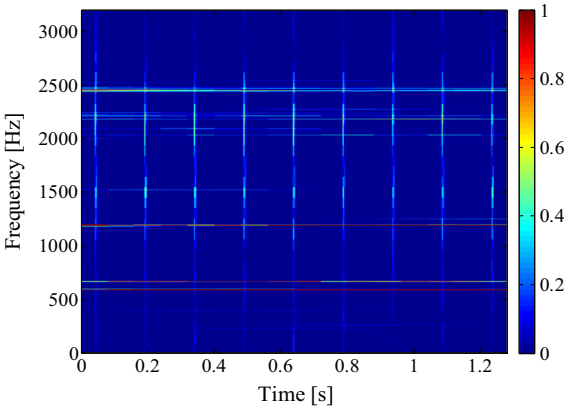

Fig. 27. The time-frequency distributions of the impulsive components.

Fig. 28. The local pit corrosion on the large gear of high-speed shaft.

the vibration signal is displayed in Fig. 23. It seems that there are some impulsive components, but the impulse period cannot be confirmed. The related frequency spectrums are illustrated in Figs. 24and 25, the first and second order of the gear meshing frequencies of the high-speed stage can be found clearly. However, there is no obvious sideband around the first order meshing frequency and the amplitude of the sideband of the second order meshing frequency is relatively low.

The above analysis fails to obtain a satisfying diagnosis conclusion. Then the SSTFA is applied to process this segment of the signal. The analysis parameter is chosen as: $l_1 = 2048$, $n_1 = 20$, $l_2 = 64$, and $n_2 = 2$. The sparse level λ_1 is chosen as 0.3 and λ_2 is chosen as 0.1 by the trial and error approach on another segment of the signals. The separated steady modulation components and impulsive components are illustrated in Fig. 26; and time-frequency distributions of different components are illustrated in Fig. 27. It can be found that there exist some period impulses with the interval of 0.153 s, which agrees with the theoretical value of the local fault in the large gear of high-speed shaft. An endoscopy examination was performed on the gearbox. A local pit corrosion and scuff were found on the large gear of high-speed shaft, as shown in Fig. 28.

5. Conclusion

Different from the classic sparse representation algorithm, the relations between the time-frequency coefficients are fully taken into consideration in the SSTFA. Theoretically, the SSTFA inherits advantages of the classic sparse representation, especially in analyzing signals with complicated time-frequency structures and extracting different feature information. Moreover, compared with the traditional time-frequency analysis methods, it is found that the SSTFA also has a better time-frequency resolution. The simulations and experiments show that the SSTFA is suitable for processing stationary speed vibration signals and has good performance for gear fault diagnosis.

The rough analysis of the parameters is given in this paper. The quantitative analysis of the parameters and self-adaptive parameters selection methods need to be further studied.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (Nos. 51405369 & 51605365), the National Key Basic Research Program of China (No. 2015CB057400), the National Natural Science Foundation of Shaanxi Province (No. 2016JQ5049), the Young Talent fund of University Association for Science and Technology in Shaanxi of China, and the Fundamental Research Funds for the Central Universities (No. xjj2014107).

References

[1] Y. Wang, F. Liu, Z. Jiang, S. He, Q. Mo, Complex variational mode decomposition for signal processing applications, Mech. Syst. Signal Process. 86 (2017) 75–85.
[2] S. Wei, Q. Han, Z. Peng, F. Chu, Dynamic analysis of wind turbine gearboxes with unknown-but-bounded parameters under random wind excitation, Iet Renewable Power Generation, 2016.
[3] R. Yan, R.X. Gao, Harmonic wavelet-based data filtering for enhanced machine defect identification, J. Sound Vib. 329 (2010) 3203–3217.
[4] R. Yan, R.X. Gao, Rotary machine health diagnosis based on empirical mode decomposition, J. Vib. Acoust. 130 (2008) 111–120.
[5] P.D. Mcfadden, A technique for calculating the time domain averages of the vibration of the individual planet gears and the sun gear in an epicyclic gearbox, J. Sound Vib. 144 (1991) 163–172.
[6] P.D. Mcfadden, Determining the location of a fatigue crack in a gear from the phase of the change in the meshing vibration, Mech. Syst. Signal Process. 2 (1988) 403–409.
[7] P.D. Mcfadden, Interpolation techniques for time domain averaging of gear vibration, Mech. Syst. Signal Process. 3 (1989) 87–97.
[8] D.S. Singh, Q. Zhao, Pseudo-fault signal assisted EMD for fault detection and isolation in rotating machines, Mech. Syst. Signal Process. 81 (2016) 202–218.
[9] J. Yuan, Z. He, Y. Zi, Gear fault detection using customized multiwavelet lifting schemes, Mech. Syst. Signal Process. 24 (2010) 1509–1528.

[10] E.C. Smith, M.S. Lewicki, Efficient auditory coding, Nature 439 (2006) 978–982.
[11] B.A. Olshausen, D.J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature 381 (1996) 607–609.
[12] Z. Feng, F. Chu, Application of atomic decomposition to gear damage detection, J. Sound Vib. 302 (2007) 138–151.
[13] Z.B. Yang, M. Radzienski, P. Kudela, W. Ostachowicz, Fourier spectral-based modal curvature analysis and its application to damage detection in beams, Mech. Syst. Signal Process. 84 (2016) 763–781.
[14] Z.B. Yang, M. Radzienski, P. Kudela, W. Ostachowicz, Two-dimensional modal curvature estimation via Fourier spectral method for damage detection, Compos. Struct. 148 (2016) 155–167.
[15] X. Chen, Z. Du, J. Li, X. Li, H. Zhang, Compressed sensing based on dictionary learning for extracting impulse components, Signal Process. 96 (2014) 94–109.
[16] W. Fan, G. Cai, Z.K. Zhu, C. Shen, W. Huang, L. Shang, Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction, Mech. Syst. Signal Process. 56–57 (2015) 230–245.
[17] L. Cui, C. Kang, H. Wang, P. Chen, Application of composite dictionary multi-atom matching in gear fault diagnosis, Sensors 11 (2011) 5981–6002.
[18] Z. Feng, M. Liang, Complex signal analysis for planetary gearbox fault diagnosis via shift invariant dictionary learning, Measurement 90 (2016) 382–395.
[19] G. He, K. Ding, H. Lin, Gearbox coupling modulation separation method based on match pursuit and correlation filtering, Mech. Syst. Signal Process. 66 (2015) 597–611.
[20] R. Jenatton, G. Obozinski, F. Bach, Structured sparse principal component analysis, J. Mach. Learning Res. 9 (2009) 131–160.
[21] R. Jenatton, J. Mairal, G. Obozinski, F. Bach, Proximal methods for sparse hierarchical dictionary learning, in: International Conference on Machine Learning, 2010, pp. 487–494.
[22] A. Lefèvre, F. Bach, C. Févotte, Itakura-Saito nonnegative matrix factorization with group sparsity, in: IEEE International Conference on Acoustics, Speech and Signal Processing, 2011, pp. 21–24.
[23] K. Siedenburg, M. Dörfler, Structured sparsity for audio signals, in: Proc. Int. Conf. on Digital Audio Effects (DAFx), 2011.
[24] B. Yang, R. Liu, X. Chen, Fault diagnosis for wind turbine generator bearing via sparse representation and shift-invariant K-SVD, IEEE Trans. Ind. Inf. (2017) (1-1).
[25] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer Publishing Company, Incorporated, 2010.
[26] D.L. Donoho, For most large underdetermined systems of linear equations the minimal L1-norm solution is also the sparsest solution, Comm. Pure Appl. Math (2006) 797–829.
[27] M. Kowalski, K. Siedenburg, M. Dorfler, Social sparsity! Neighborhood systems enrich structured shrinkage operators, IEEE Trans. Signal Process. 61 (2013) 2498–2511.
[28] M. Kowalski, B. Torrésani, Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients, SIViP 3 (2009) 251–264.
[29] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Commun. Pure Appl. Math. 57 (2004) 1413–1457.
[30] S. Kai, M. Doerfler, Persistent time-frequency shrinkage for audio denoising, J. Audio Eng. Soc. 61 (1/2) (2013) 29–38.
[31] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, Siam J. Imaging Sci. 2 (2009) 183–202.
[32] M. Kowalski, Sparse regression using mixed norms, Appl. Comput. Harmonic Anal. 27 (2009) 303–324.
[33] Y. Qin, Y. Mao, B. Tang, Vibration signal component separation by iteratively using basis pursuit and its application in mechanical fault detection, J. Sound Vib. 332 (2013) 5217–5235.
[34] Z. Feng, Y. Zhou, M.J. Zuo, F. Chu, X. Chen, Atomic decomposition and sparse representation for complex signal analysis in machinery fault diagnosis: a review with examples, Measurement 103 (2017) 106–132.
[35] L.C. Freudinger, R. Lind, M.J. Brenner, Correlation filtering of modal dynamics using the Laplace wavelet, in: Proceedings of SPIE - The International Society for Optical Engineering, vol. 1999, 1998.
[36] J.L. Starck, Y. Moudden, J. Bobin, M. Elad, D.L. Donoho, Morphological component analysis, Opt. Photon. (2005) 31–41.
[37] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N. 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. Lond. A: Math. Phys. Eng. Sci. 454 (1998) 903–995.