Using Wavelet Support Vector Machine for Fault Diagnosis of Gearboxes

Mohammad Heidari[1]

Department of Mechanical Engineering, Aligudarz Branch, Islamic Azad University, Aligudarz, Iran

Abstract

Identifying fault categories, especially for compound faults, is a challenging task in mechanical fault diagnosis. For this task, this paper proposes a novel intelligent method based on wavelet packet transform (WPT) and multiple classifier fusion. An unexpected damage on the gearbox may break the whole transmission line down. It is therefore crucial for engineers and researchers to monitor the health condition of the gearbox in a timely manner to eliminate the impending faults. However, useful fault detection information is often submerged in heavy background noise. The non-stationary vibration signals were analyzed to reveal the operation state of the gearbox. The proposed method is applied to the fault diagnosis of gears and bearings in the gearbox. The diagnosis results show that the proposed method is able to reliably identify the different fault categories which include both single fault and compound faults, which has a better classification performance compared to any one of the individual classifiers. The vibration dataset is used from a test rig in Shahrekord University and a gearbox from Sepahan Cement. Eventually, the gearbox faults are classified using these statistical features as input to WSVM.

Keywords*: gearbox, fault diagnosis, wavelet, support vector machine*

[DOI: 10.22068/ijae.8.1.2603]

1. Introduction

Gearboxes are widely employed in a variety of industrial applications. The breakdowns of the transmission machinery resulted from the gearbox failures account for 80% and the malfunctions of the gearbox are mostly caused by the gear and bearing faults [1]. Therefore, since the efficient running of machinery plays an influential part in the economics of an organization, it is necessary and critical that the early detection and diagnosis of gear and bearing faults should be performed to prevent breakdown accidents and reduce the economic loss. Up to date great importance has been paid to the studies aiming at gearbox-fault identification and characterization. Among available diagnosis techniques vibration analysis is manifestly the most commonly used and also very efficient method because vibration signals can capture important dynamic information that reflects the states of the machines [1–4]. Classical techniques, including the spectral analysis [5], time domain averaging [6], and time-series analysis [7], have been proven to be effective for the stationary vibration signals. However, in reality the measured data is inherently nonstationary. In order to analyze nonstationary signals, a series of advanced techniques have been developed, such as Wigner–Ville distributions (WVDs) [8], empirical mode decomposition (EMD) [9], S transform [10], order tracking technique [11] and wavelet transform (WT) [12], etc. However, in practice the vibration signal acquisition always involves with a multi-channel sensor system and these mentioned methods have been widely used to process a single channel sensor data. Thus it is inevitable to assess the efficiency of all the sensors prior to the diagnosis to choose an optimal channel [2]. However, the background noise makes it difficult in determining the optimal channel [2]. To overcome this problem, the independent component analysis (ICA) is developed to find a suitable representation of multi-channel sensors, which is called the blind source separation (BSS) [13]. Each sensor data of the gearbox can be regarded as the mixture of different vibration components, including the gear meshing vibration, body vibration of the gearbox, vibration of the rolling bearings, and the noise signal, etc. ICA can be useful in investigations of vibration signals, by separating out noise contributions from mixed sources and focusing on the correlated

characteristics which presumably comes from the gear faults [14]. Lin et al. [15,16] combined the ICA with WT to detect the gear failures. Both the numerical simulation and experimental tests show that the ICA can efficiently extract independent features of the gear meshing and simultaneously reduce the effect of noise. Yang et al. [17] proved that by means of ICA and the further feature extraction strategy based on residual mutual information (RMI), higher than second order features embedded in multi-channel vibration measurements can be captured effectively. Literature [18–20] also had shown the usefulness of the ICA for the gear fault detection. Nevertheless, in their work the noise sources were supposed to be mixed linearly with the gear fault vibration and the linear ICA was employed to process the experimental vibration data. Since the gearbox vibration signals obtained from experiment/practice are usually the nonlinear mixture, the performance of the linear ICA still has much room for improvement in these applications. The kernel ICA (KICA) proposed by Bach and Jordan [21] focuses on the nonlinear BSS problems and can separate the desired source from the nonlinear mixed noise. Tian et al. [22] used the KICA to process the transient acoustic signal on gearbox, the analysis results showed that the KICA is more robust than ICA, and the fault characters were fully detected. The validity of the approach and outcomes should be tested for the vibration signal of the gearbox. As widely recognized, one challenging task in fault diagnosis is feature extraction and selection. In general, the original fault characteristics obtained from advanced signal processing methods (i.e. WVD, EMD and WT, etc.) contain some redundant information. If use them as the signatures to assess the gearbox health state, the performance may be unsatisfactory. Hence, it is essential to remove the useless features. The problem is that it is not easy to determine the distinct features. One of the most popular methods, principal component analysis (PCA) has been proven to be effective for feature reduction [23]. However, its main limitation lies in the ability of preserving the nonlinear properties of the original feature space [24–26]. Fortunately, the manifold learning algorithms provide a new means to dealing with the nonlinear dimensionality reduction problems. Compared with the linear methods, the purpose of manifold learning is to project the original high-dimensional data into a lower dimensional feature space by preserving the local topology of the original data. Thus, the intrinsic structure of the data of interest can be extracted effectively. Successful applications of these new nonlinear feature selection methods can be found in the field of image processing, speech spectrograms, electroencephalography (EEG) and electrocardiograph (ECG) signals for medical diagnose [26]. Furthermore, very limited work has been done to address the multiple faults detection of gearboxes using manifold learning algorithm. Yang et al. [27] proposed a method for nonlinear time series noise reduction based on principal manifold learning applied to the analysis of gearbox vibration signal with tooth broken, but only for signal denoising. Li et al. [28] proposed the multiple manifolds analysis (MMA) approach to extract manifold information from the bearing vibration signals with different faults and Wang et al. [29] combined locally linear embedding (LLE) and kernel fisher discriminant analysis (KFDA) to detect rolling bearing fault. In the previous work we also adopted the LLE algorithm for the feature reduction of the gear crack level identification [3]. However, the nonlinear BSS problem was not considered in these studies. Hence, it is worth investigating the fault diagnosis performance for both single and coupled faults of the gearbox

(including the gears and rolling bearings) by using the integration of the KICA and manifold learning. This paper aims to deal with multi-fault diagnosis of the gearbox, including gear faults and rolling bearing faults. A method is proposed based on the KICA, LLE and fuzzy knearest neighbor (FKNN). In comparison with the fault diagnosis method based on manifold learning reported in [3,28,29], the proposed technique in this work adopts not only nonlinear dimensionality reduction algorithm, but also KICA for nonlinear BSS problem. Thus, it possesses a more powerful fault diagnosis capability than existing approaches. The effectiveness of the proposed approach is demonstrated by two case studies. Xian and Zeng [30] developed an intelligent fault diagnosis procedure based on wavelet packet transform (WPT) and hybrid SVM. Zamanian and Ohadi [31] presented a method for feature extraction based on exact wavelet analysis to improve the fault diagnosis of gears. In their study, feature extraction was based on maximization of local Gaussian correlation function of wavelet coefficients. They used from a linear support vector machine to classify feature sets extracted with the presented method. The rest of this paper is outlined as follows. Section 2 briefly describes the fundamental theory of wavelet packet decomposition and two wavelet selection criteria. The proposed new machine health status identification method is presented in Section 3, followed by the experimental verification tests using both bearing and gearbox datasets as stated in Section 4. In Section 5, the effect of different wavelet basis functions on the performance of the proposed scheme is discussed. Conclusions are drawn in Section 6.

[DOI: 10.22068/ijae.8.1.2603]

2. Theoretical background

2.1. The review of wavelet packet transform

The wavelet method is a signal processing technique to represent and analyse a time signal in the time-frequency domain. This method is based on the shifted and scaled signal decomposition of a prototype function called mother wavelet [27,31]. These functions are similar to the complex sinusoid used in the Fourier transform, except for two fundamental differences: (1) the complex sinusoid lasts infinitely, whereas the wavelets are functions of limited duration, which are located in time (translation) and frequency (dilatations); and (2) the sinusoid is smooth and predictable, whereas the wavelet tends to be irregular and asymmetric.

Let $\psi(t) \in L^2(R)$ be a function called mother wavelet, then $\psi_{s,u}(t)$, with s, u ∈ r, and s>0 are a family of shifted and scaled functions of a mother wavelet. This provides a modulated window $\psi(t)$, which generates an entire family of elementary functions by dilatations or contractions, and translations in time defined by Eq. (1) [27,31]:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi(\frac{t-u}{s})$$

Where s is the scaling parameter, and u the position parameter. The wavelet transform (WT) in continues time of a function is called a continuous wavelet transform (CWT), which is calculated by the inner product of the analysed signal with a family of shifted and scaled wavelets, using the expression Eq. (2) [27,31]:

$$CWTx(s,u) = < x(t), \psi_{u,s}(t) > = \frac{1}{\sqrt{s}}\int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t-u}{s}\right)dt$$

The CWT is a useful method for the analysis of non-stationary signals with different behaviours during the sampling time, enabling the temporal location of the components in the frequency domain. The main drawback is low analytical-computational efficiency limiting its use to off-line applications. For applications requiring real-time signal processing, wavelet analysis is performed using two methods termed discrete wavelet transform (DWT), and wavelet packet transform (WPT). Both methods decompose the signal into a mutually orthogonal set of wavelets, derived from the application of a pyramidal algorithm of convolutions with quadrature mirror filters, based on the coefficients described in Eqs. (3) and (4)

$$A_j(k) = \sum_n h(n-2k)c_{j-1}(n)$$

$$D_j(k) = \sum_n g(n-2k)c_{j-1}(n)$$

Where A_j (k) and D_j (k) are scaling and wavelet coefficients, j is the number of transformation levels with j=1, 2, ...; k is the number of scaled and wavelet coefficients with k=1, 2, ..., N×2-j

, where N is the total number of samples of the original signal; h and g are low-pass and high-pass coefficients of the scaled function and wavelet function, respectively, based on a chosen mother wavelet; and n is the filter length. These coefficients successively decompose the original signal into approximation (low frequency) or detail (high frequency) signals using the scaled and wavelet coefficients, respectively. In the WPT method for 3-level decomposition, at level L1 (j=1) the original signal is decomposed in two frequency ranges: an approximation A (scaling coefficients) is calculated using a low-pass filter (H), and a detail D (wavelet coefficients) calculated with a high-pass filter (G). Low-pass filters remove high frequency fluctuations and preserve slow trends, and high-pass filters remove the slow trends and preserve high frequency. After filtering, the original signal is decimated by a factor of 2, so that the approximation and detail coefficients are equal in number to the sample data of the original signal. Moreover, this procedure eliminates redundant information and significantly enhances the performance of the algorithm. At decomposition level L2 (j=2), A and D are subdivided into approximation AA and AD, and detail DA and DD coefficients, respectively. At level L3 (j=3) the procedure is repeated. The approximation and detail coefficients generate at each level independent frequency packets consisting of Nx2-j coefficients. This procedure is repeated until the desired wavelet decomposition level is achieved. In this paper optimal decomposition level of wavelet is selected based on the Maximum Energy to Shannon Entropy ratio criteria [32].

3. Multi class wavelet SVM

In this section, the wavelet kernel, OAA and OAO MSVM strategies are presented.

3.1. Wavelet kernel

[DOI: 10.22068/ijae.8.1.2603]

The WSVM aims at finding the ideal classification in the space spanned by multidimensional wavelet. The concept behind the wavelet analysis is to express a signal by a family of functions generated by h(x) called mother wavelet [33]:

$$h_{a,c}(x) = |a|^{-1/2} h\left(\frac{x-c}{a}\right)$$

where x, a, c ∈ R, a is a dilation factor, and c is a translation factor. A common multidimensional wavelet function can be expressed as the product of 1D wavelet function [34]:

$$h(x) = \prod_{i=1}^{N} h(x_i)$$

where N is the dimension number. Let h(x) denotes a mother kernel function. Then dot product wavelet kernels are:

$$k(x, \acute{x}) = \prod_{i=1}^{N} h(\frac{x_i - c_i}{a}) h\left(\frac{\acute{x}_i - \acute{c}_i}{a}\right)$$

The decision function for classification is:

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i \prod_{j=1}^{N} \psi(\frac{x^i - x_i^j}{a_i})\right) + b$$

In this paper, four kernel functions are used: wavelet Morlet, wavelet Maxican hat, Gaussian wavelet kernel and wavelet Shannon. The multi-class classification strategy, such as OAA, OAO and OAOT with different wavelet kernel functions is used for classification in this paper [35].

4. Experimental setup

4.1. Case study 1: Shahrekord test rig

The experimental setup at Shahrekord University to collect dataset consists of a one-stage gearbox with spur gears, a flywheel and an electrical motor. The test rig has been shown in Figure 1. Vibration signals are obtained in the radial direction by mounting the accelerometer on the top of the gearbox. "Easy Viber" data collector and its software, "SpectraPro", are used for data acquisition. The sensitivity and dynamic range of accelerometer probe are 100mv/g and ±50 g. The signals are sampled at 16000 Hz lasting 2 s. In the present study, four pinion wheels are used. The vibration signal from accelerometer is captured for the following conditions: good gear, gear with tooth breakage, chipped tooth gear and eccentric gear. For bearing vibration signal acquisition four self-aligning ball bearings (1209K) are used. One new bearing is considered as good bearing. In the other three bearings, some defects are created and then various bearings are installed and the raw vibration signals acquired on the bearing housing. So the vibration signals are captured for the following conditions: good bearing, bearing with spall on inner race, bearing with spall on outer race, bearing with spall on ball and bearing with combine defect. Figure 2 shows the faults in the baering of Shahrekord test rig.

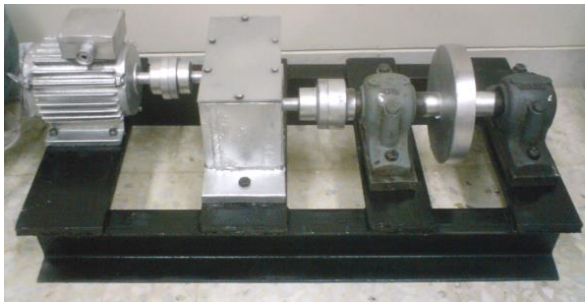

Fig1.. Fault simulator set up in Shahrekord University [36].

Fig2. Bearing component with fault (Shahrekord test rig), outer race fault, inner race fault and ball fault

[DOI: 10.22068/ijae.8.1.2603]

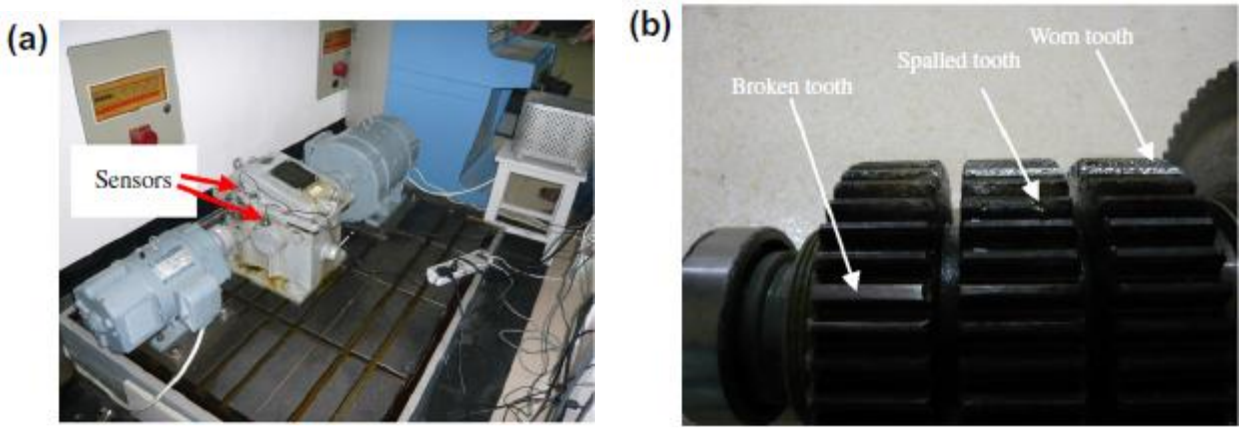

Fig3. (a) The gear fault test setup (Sepahan Cement), and (b) the appearance of the faulty gears.

4.2. Case study 2: gear fault diagnosis

The gear fault simulator setup with sensor is shown in Figure 3. It is a two-stage gear transmission system and its transmission path is as:

Input ⟶ (Z26/Z64) ⟶ (Z40/Z85)⟶ output.

Two piezoelectric accelerometers (CA-YD-106) are mounted on the flat surface of the input shaft and output shaft to collect the gearbox vibrations. The vibrations have been measured under six different gear conditions: pattern A-normal, pattern B-single cracked tooth, pattern C-single spalled gear, pattern D-single broken tooth, pattern E-compound fault of cracked and broken tooth, pattern F-compound fault of worn and spalled tooth. The single gear fault to be analyzed was introduced on the 40-tooth gear carried on the transmission shaft, and the compound fault occurred in the Z40/Z85 gear pair. The vibration data is acquired under 750 rpm with heavy-load. The sampling frequency was 10,000 Hz and sample length was 19,456 for all conditions. Each sensor has sampled 50 times for every gear condition and the total samples are 300 for each sensor.

5. Result and discussion

Six base wavelets such as Meyer, symlet16, cofi5, rbio6.8, bior6.8 and db44 are selected for this work. Based on two wavelet selection cratria, Daubechies wavelet (db44) and Meyer are selected as the best base wavelet among the other wavelets considered from the Maximum Relative Energy and Maximum Energy to Shannon Entropy criteria respectively [37]. The wavelet packet coefficients of all signals with db44 and Meyer are calculated at the eighth level of decomposition. After WPT, 2304 statistical features are extracted from the 256 nodes at eight decomposition levels. When applying wavelet transform to a signal, if the Shannon entropy measure of a particular scale is minimum then we can say that a major defect frequency component exists in the scale but, in the present study out of 256 scales considered, the scale having the Maximum Energy to Shannon Entropy of healthy condition is selected, and the statistical features of the wavelet packet coefficient corresponding to the selected level are calculated. Statistical moments like kurtosis, skewness and standard deviation are descriptors of the shape of the amplitude distribution of vibration data, and have some advantages over traditional time and frequency analysis, such as its lower sensitivity to the variations of load and speed. In the present paper, authors' use statistical moments like standard deviation, crest factor, absolute mean amplitude value, variance, kurtosis, skewness and fourth central moment as features to effectively indicate early faults occurring in rolling element bearings and gears. In addition, energy and Shannon entropy of the wavelet coefficients are used as two new features along with other statistical parameters as input of the classifier. These statistical features are fed as input to the soft computing techniques like SVM for fault classification. Two cases of input data and feature sets are considered for classification. In case A, statistical parameters of wavelet packet transform are considered (for each type of the gearbox fault). Case B is related to the condition that statistical features in optimal level, which has been extracted based on the criteria of Maximum Energy to Shannon Entropy ratio, are considered (for each type of gearbox fault). In addition, energy and Shannon

[DOI: 10.22068/ijae.8.1.2603]

entropy factors are used as two new features as features sets in this case. Table 1 shows the results of classification of gearbox with Maximum Energy to Shannon Entropy criterion. In the case B, by Maximum Energy to Shannon Entropy ratio criterion (Table 1), for test set, correctly classified instances is 96.25%. While using 10-fold cross validation average classification accuracy is 95.41%.

Table 2 shows accuracy for fault classification with Maximum Relative Wavelet Energy criterion. The correctly classified (case B) instances using test set is 92.91% . For 10-fold cross validation, average classification accuracies for WSVM is 92.08%. which is slightly less than the previous case. From Tables 1 and 2, we found that the Maximum Energy to Shannon Entropy criterion with two new features is better for fault classification of gearbox with respect to Maximum Relative Wavelet Energy criterion. Furthermore, the accuracy comparison of WSVM with OAOT, OAA and OAO with Maximum Energy to Shannon Entropy is listed in Table 3. From Table 3, it is clear the proposed method based on wavelet support vector machine using the Morlet wavelet kernel has improved the classification accuracy by 8.09% with respect to Shannon wavelet kernel. In this case, the overall average classification accuracy is 95%. From Table 3, we find that the classification accuracy with OAOT strategy is better than OAA and OAO.

Figures 4 shows the testing time and training time of WSVM with three strategies. We can observe that the training time in OAA is bigger than in OAO and OAOT under all kernel functions. As shown in Figure 4, the performance of the Morlet kernel for machinery fault diagnosis is acceptable. From Figure 4, we find that the Morlet kernel has the least testing and training time with respect to other kernel functions. It is clear from Figure 5, the multi kernel has the least training and testing time with OAOT algorithm. Therefore the OAOT strategy is better than OAO and OAA for the problem.

Figure 5 shows that the accuracy of WSVM using OAOT algorithm with Mexican hat kernel reaches the highest point (94.16%) with C=38.7 and a=0.83. Similarly, when we apply the Mexican hat kernel to OAO algorithm and OAA algorithm, the best classification ratio is 86.24% and 90.41%, respectively. Figure 4 shows that the accuracy of WSVM using OAOT algorithm with the Morlet kernel function reaches the highest point (96.24%) with C=29.7 and a=0.74. Similarly, when we apply the Morlet kernel to OAO algorithm and OAA algorithm, the best classification ratio with same a, and C is 89.16% and 92.91%, respectively. Figure 5 shows that the accuracy of MSVM using OAOT algorithm with the Shannon kernel reaches the highest point (86.66%) with C=50 and a=0.4. Similarly, when we apply the Shannon kernel to OAO algorithm and OAA algorithm, the best classification ratio is 81.66% and 84.16%, respectively.

Figure 8 shows that the accuracy of MSVM using OAOT algorithm with the Gaussian kernel reaches the highest point (91.66%) with C=100 and a=0.5. Also, when we apply the Gaussian kernel to OAO algorithm and OAA algorithm, the best classification ratio is 85.41% and 88.74%, respectively.

Table 1. Classification performance (Maximum Energy to Shannon Entropy criterion)

Parameters		WSVM (with Morlet kernel)	
		Test set	10-fold cross validation
Correctly classified	Case A	225 (93.75%)	222 (92.50%)
	Case B	231 (96.25%)	229 (95.41%)
Incorrectly classified	Case A	15 (6.25%)	18 (7.5%)
	Case B	9 (3.75%)	11 (4.85%)
Total number of instances		240	240
Training time (s)	Case A (WSVM)	137.41	
	Case B (WSVM)	140.90	

[DOI: 10.22068/ijae.8.1.2603]

Table 2. Classification performance (Maximum Relative Wavelet Energy criterion)

Parameters		WSVM (with Morlet kernel)	
		Test set	10-fold cross validation
Correctly classified	Case A	218 (90.83%)	215 (89.58%)
	Case B	223 (92.91%)	221 (92.08%)
Incorrectly classified	Case A	22 (9.16%)	25 (10.41%)
	Case B	17 (7.08%)	19 (7.91%)
Total number of instances		240	240
Training time (s)	Case A (WSVM)	144.28	
	Case B (WSVM)	149.05	

Table 3. The classified result of experiment data using WSVM with three methods

Operating condition		Fault classification accuracy based on SVM with kernel (%)			
		Morlet *C=29.7, a=0.74*	Mexican hat *C=38.7, a=0.83*	Gaussian	Shannon
Outer race fault	OAOT	96.66	93.33	90	86.66
	OAA	93.33	93.33	90	83.33
	OAO	90	86.33	86.33	83.33
Inner race fault	OAOT	96.66	96.66	93.33	90
	OAA	93.33	90	86.66	83.33
	OAO	90	86.66	86.66	83.33
Roller fault	OAOT	100	96.66	93.33	86.66
	OAA	96.66	93.33	90	83.33
	OAO	93.33	90	90	83.33
Combine fault	OAOT	96.66	93.33	93.33	86.66
	OAA	93.33	90	90	83.33
	OAO	90	86.66	83.33	80
Average accuracy (bearing)	OAOT	97.49	94.99	92.49	86.66
	OAA	94.16	91.66	89.16	83.33
	OAO	90.83	87.49	86.66	82.49
Chipped tooth gear	OAOT	100	96.66	96.66	86.66
	OAA	96.66	90	90	86.66
	OAO	86.66	83.33	83.33	80
Eccentric gear	OAOT	93.33	93.33	90	86.66
	OAA	90	86.66	86.66	83.33
	OAO	90	83.33	83.33	80
Broken-tooth gear	OAOT	93.33	90	86.66	83.33
	OAA	90	90	86.66	83.33
	OAO	86.66	86.66	83.33	80
Good gearbox	OAOT	93.33	93.33	90	90
	OAA	90	90	90	86.66
	OAO	86.66	86.66	86.66	83.33
Average accuracy (gear)	OAOT	94.99	93.33	90.83	86.66
	OAA	91.66	89.16	88.33	84.99
	OAO	87.49	84.99	84.16	80.83

[DOI: 10.22068/ijae.8.1.2603]

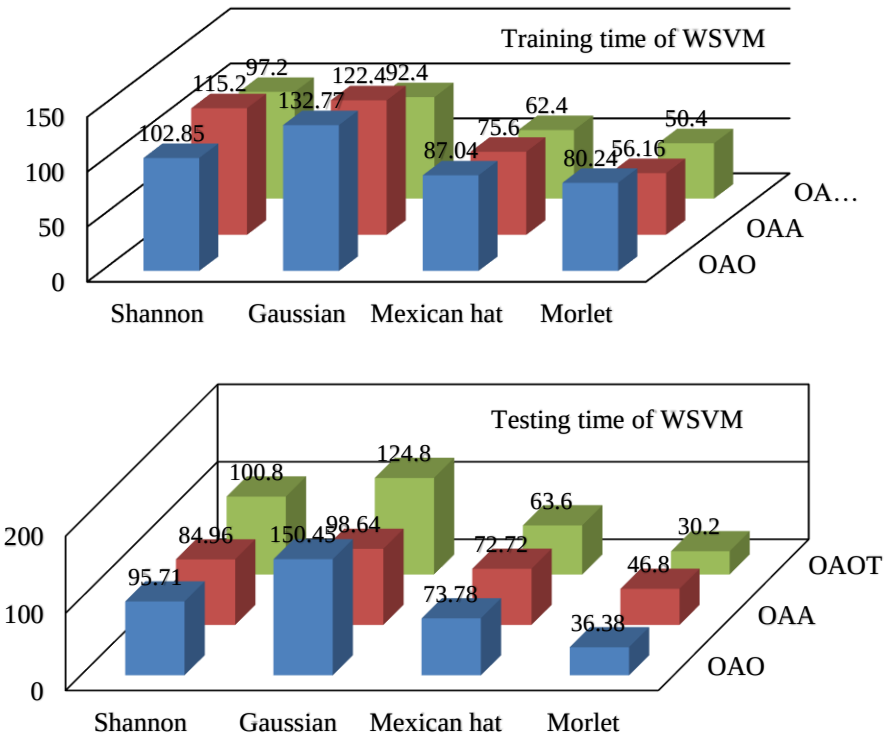

Fig4. Training time and testing time for WSVM

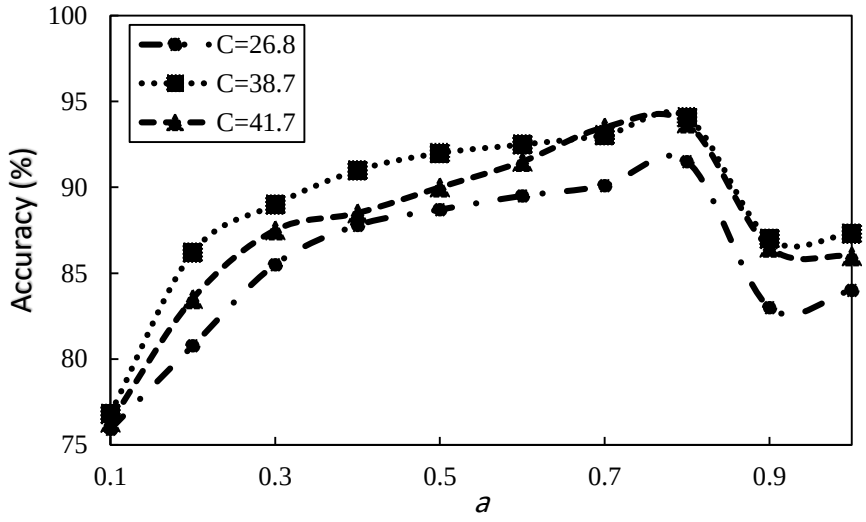

Fig5. Comparison of accuracy using OAOT algorithm based on WPT feature extraction with Mexican hat kernel in different (C, a)

[DOI: 10.22068/ijae.8.1.2603]

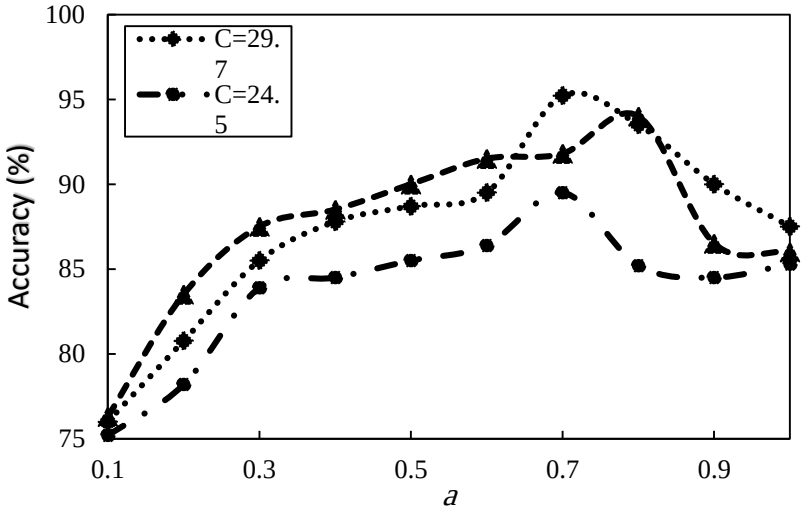

Fig6. Comparison of accuracy using OAOT algorithm based on WPT feature extraction with Morlet kernel in different (C, a)

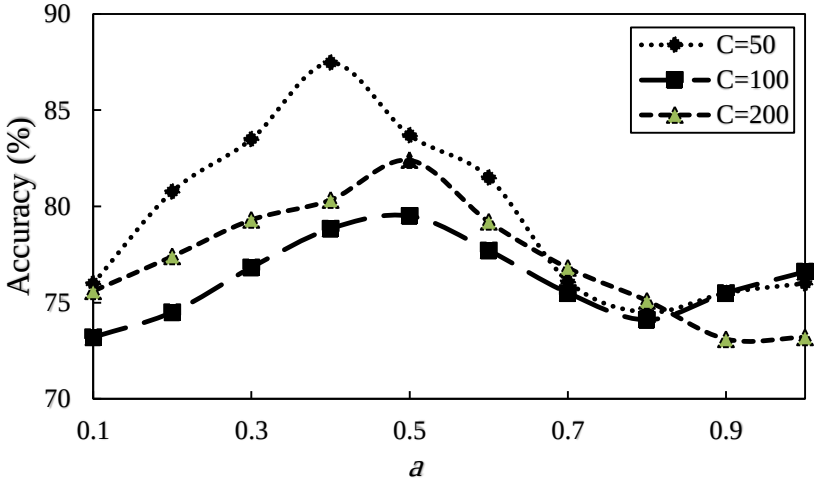

Fig7. Comparison of accuracy using OAOT algorithm based on WPT feature extraction with Shannon kernel in different (C, a)

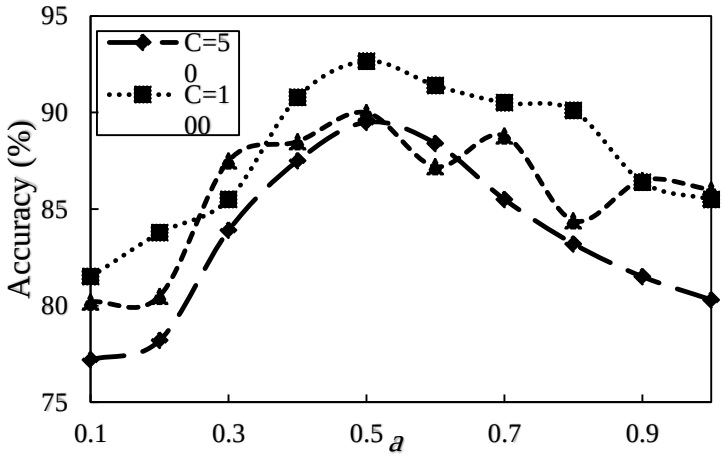

Fig8. Comparison of accuracy using OAOT algorithm based on WPT feature extraction with Gaussian kernel in different (C,a)

[DOI: 10.22068/ijae.8.1.2603]

6. Conclusions

Impending fault detection and identification for gearboxes, (specifically, the gears and rolling bearings), is of great importance to the reliable operation throughout their service lives. The use of the vibration signals is a promising way to assess the health condition of the gearbox in practice. However, the recorded vibrations are always corrupted by heavy background noise. Therefore, a new diagnostic method combining the wavelet transform and wavelet support vector machine has been proposed for the gearbox multi-fault diagnosis in this work. The experimental tests on the gear failures and rolling bearing faults have been implemented and presented to verify the efficacy of the newly proposed approach. The two experimental case studies show that (a) the experimental vibration signals on the gearbox can be demixed by the WPT with a small amount of information losing, (b) nonlinear property of the fault characteristics can be captured by the WSAVM, and hence the fault detection rate has been increased, and (c) the proposed intelligent diagnosis method can provide satisfactory fault detection performance for the gearbox. The newly proposed method can enhance the ability of fault detection for both the gears and rolling bearings. Thus, the proposed diagnosis approach in this work may provide practical and effective utilities for gearbox fault diagnosis as well as other complex machines such as vehicle transmission systems in future research.

References

[1]. Z. Li, X. Yan, C. Yuan, Z. Peng, L. Li, Virtual prototype and experimental research on gear multi-fault diagnosis using waveletautoregressive model and principal component analysis method, Mechanical Systems and Signal Processing, 25 (2011) 2589–2607.

[2]. A. Albarbar, F. Gu, A. Ball, Diesel engine fuel injection monitoring using acoustic measurements and independent component analysis, Measurement, 45 (2010) 1376–1386.

[3]. Z. Li, X. Yan, Y. Jiang, L. Qin, J. Wu, A new data mining approach for gear crack level identification based on manifold learning, Mechanika, 18 (2012) 29–34.

[4]. Z. Shen, X. Chen, X. Zhang, A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM, Measurement, 45 (2012) 30–40.

[5]. J. Sanza, R. Pererab, C. Huertab, Fault diagnosis of rotating machinery based on auto-associative neural networks and wavelet transforms, Journal of Sound and Vibration, 302 (2007) 981–999.

[6]. P.D. McFadden, Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration, Mechanical Systems and Signal Processing, 1(1987) 173–183.

[7]. C. Wang, Y. Kang, P. Shen, Y. Chang, Y. Chung, Applications of fault diagnosis in rotating machinery by using time series analysis with neural network, Expert Systems with Applications, 37 (2010) 1696–1702.

[8]. N. Baydar, A. Ball, A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distributions, Mechanical Systems and Signal Processing, 15 (2001) 1091–1107.

[9]. Y. Yang, Y. He, J. Cheng, A gear fault diagnosis using Hilbert spectrum based on MODWPT and a comparison with EMD approach, Measurement, 42 (2009) 542–551.

[10]. B. Li, P. Zhang, H. Tian, S. Mi, D. Liu, G. Ren, A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox, Expert Systems with Applications, 38 (2011) 10000–10009.

[11]. M. Bai, J. Huang, M. Hong, F. Su, Fault diagnosis of rotating machinery using an intelligent order tracking system, Journal of Sound and Vibration, 280 (2005) 699–718.

[12]. S. Wang, W. Huang, Z. Zhu, Transient modeling and parameter identification based on wavelet and correlation filtering for rotating machine fault diagnosis, Mechanical Systems and Signal Processing, 25 (2011) 1299–1320.

[13]. A. Hyvärinen, Survey on independent component analysis, Neural Computing Surveys, 2 (1999) 94–128.

[14]. M.J. Roan, J.G. Erling, L.H. Sibul, A new, non-linear, adptive, blind source separation approach to gear tooth failure detection and analysis, Mechanical Systems and Signal Processing, 16 (2002) 719–740.

[15]. J. Lin, A. Zhang, Fault feature separation using wavelet-ICA filter, NDT and E International, 38 (2005) 421–427.

[16]. M.J. Zuo, J. Lin, X. Fan, Feature separation using ICA for a one dimensional time series and its application in fault detection, Journal of Sound and Vibration, 287 (2005) 614–624.

[17]. S. Yang, W. Jiao, Z. Wu, Multi-neural networks for faults diagnosis based on ICA feature

[DOI: 10.22068/ijae.8.1.2603]

extraction, Journal of Vibration Engineering, 17 (2004) 438–442.
[18]. Q. He, Z. Feng, F. Kong, Detection of signal transients using independent component analysis and its application in gearbox condition monitoring, Mechanical Systems and Signal Processing, 21(2007) 2056–2071.
[19]. J. Huang, S. Bi, H. Pan, X. Yang, Application of independent component analysis to fault diagnosis of gearbox, Journal of Vibration, Measurement and Diagnosis, 28 (2008) 126–130.
[20]. Z. Li, X. Yan, C. Yuan, L. Li, Gear multi-faults diagnosis of a rotating machinery based on independent component analysis and fuzzy k nearest neighbor, Advanced Materials Research,108–111 (2010)1033–1038.
[21]. F.R. Bach, M.I. Jordan, Kernel independent component analysis, Journal of Machine Learning Research, 3 (2003) 1–48.
[22]. H. Tian, L. Tang, G. Tian, Y. Zhang, Fault diagnosis of gearbox based on KICA, Journal of Vibration and Shock, 28 (2009) 163–164.
[23]. A. Widodo, B. Yang, Wavelet support vector machine for induction machine fault diagnosis based on transient current signal, Expert Systems with Applications, 35 (2008) 307–316.
[24]. M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural Computation, 15 (2003) 1373–1396.
[25]. J.B. Tenenbaum, V. Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science, 290 (2000) 2319–2323.
[26]. S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science, 290 (2000) 2323–2326.
[27]. J. Yang, J. Xu, D. Yang, Noise reduction method for nonlinear time series based on principal manifold learning and its application to fault diagnosis, Chinese Journal of Mechanical Engineering (English Edition), 42 (2006) 154–158.
[28]. M. Li, J. Xu, J. Yang, D. Yang, D. Wang, Multiple manifolds analysis and its application to fault diagnosis, Mechanical Systems and Signal Processing, 23 (2009) 2500–2509.
[29]. G. Wang, Y. Liu, X. Zhao, Fault diagnosis of rolling bearings based on LLE–KFDA, Material Science Forum, 626 (2009) (2009) 529–534.
[30]. G.M. Xian, B.Q. Zeng, An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines, Expert Systems with Applications, 36(2009) 12131–12136.
[31]. A.H. Zamanian, A. Ohadi, Gear fault diagnosis based on Gaussian correlation of vibrations signals and wavelet coefficients, Applied Soft Computing, 11(2011) 4807–4819.
[32]. K. Zhu, Y.S. Wong, G.S. Hong, Wavelet analysis of sensor signal for tool condition monitoring: a review and some new results, Int. J. Mach. Tool Manuf.49 (2009) 537–553.
[33]. I. Daubechies, Ten Lectures on Wavelets, Philadelphia, SIAM, PA, 1992.
[34]. R.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Trans. Inf. Theory, 38 (2) (1992) 713–718.
[35]. M. Heidari, S. Stanford, Wavelet support vector machine and multi-layer perceptron neural network with continues wavelet transform for fault diagnosis of gearboxes, Journal of Vibroengineering, 19(2017) 125-137.
[36]. M. Heidari, H. Homaei, H. Golestanian, A. Heidari, Using PCA with LVQ, RBF, MLP, SOM and continuous wavelet transform for fault diagnosis of gearboxes, IJAE. 5(2) (2015)1039-1053.
[37]. M. Heidari, H. Homaei, H. Golestanian, Fault diagnosis of gearboxes using LSSVM and WPT, IJAE, 5(4)(2015) 2017-2027.

[DOI: 10.22068/ijae.8.1.2603]